Topologically Complex Molecules: Synthesis and Properties

Gianga, Tiberiu-Marius

Award date:
2020

Awarding institution:
University of Bath

Topologically Complex Molecules: Synthesis and Properties

Tiberiu-Marius Gianga

A thesis submitted for the degree of Doctor of Philosophy

University of Bath

Department of Chemistry

September 2019

COPYRIGHT

Declaration of any previous submission of the work

The material presented here for examination for the award of a higher degree by research has not been incorporated into a submission for another degree.

Declaration of authorship

I am the author of this thesis, and the work described therein was carried out by myself, with the exception of parts from Chapter 2, 3 and 5 (as detailed in the thesis) where of the work was carried out by other researchers (*i.e.* Master/visiting students working under my direct supervision).

Abstract

The first chapter outlines some of the latest advancements in the field of dynamic combinatorial chemistry (DCC) and highlights different interlocked molecules synthesised using this approach.

The second chapter describes the design and synthesis of a new π-donor building block for DCC (BDT – benzo[1,2-*b*:4,5-*b'*]dithiophene derivative) and its behaviour in the presence of an acceptor molecule (NDI – naphthalenediimide) in aqueous system. The most peculiar feature of this building block is the difference in its reactivity towards the enantiomers of NDI. This means different dynamic combinatorial library (DCL) distributions are obtained depending on the NDI enantiomer used. In each case, a different type of [2]catenane is formed, giving rise to the first ever structural divergent synthesis using DCC. This is studied and discussed in detail as well as the outcome of the 1:1 enantiomeric mixture NDI mixture in the presence of BDT. This is compared with the libraries containing the 1:1 enantiomeric mixture of BDT and either enantiomer of NDI. Lastly, a racemic mixture of the another π-donor building block (DN – dialkoxynaphthalene – also extensively studied in Sanders' / Pantoș' groups) in the presence of NDI is discussed.

The third chapter reports the discovery of an aromatic all-donor [2]catenane, and the synthesis of the structural isomer of BDT (iso-BDT – benzo[1,2-*b*:5,4-*b'*]dithiophene). BDT on its own in the presence of salt (which increases the hydrophobic effect) leads to the first ever aromatic all-donor catenane formed through DCC. Its isomer, iso-BDT, does not form any catenane either on its own, or in mixed libraries with NDI (only macrocycles are obtained).

The fourth chapter introduces the design of a new homochiral acceptor-donor-acceptor trimer. The library of this trimer on its own in the presence of salt contains just one major species. The analysis of this shows a very complex ^{1}H NMR spectrum; NMR, MS and computational studies suggest that this is a Figure-of-Eight knot. When the library concentration is decreased ten times, a [2]catenane is formed instead of the knot. Very interesting is the chirality mismatch: the homochiral trimer forms a knot, while the heterochiral one leads to a [2]catenane.

In the next chapter (Chapter 5), the design and behaviour in DCLs of another π-donor molecule (BDF – benzo[1,2-*b*:4,5-*b'*]difuran derivative) is reported. In

comparison to BDT, this one does not form an all-donor catenane, but the structural divergence is present again and it is more clean than for BDT.

Chapter 6 gives a brief conclusion of all work reported in this thesis and suggests directions for future work.

Chapter 7 is the Experimental Section, including all synthetic routes used in this work, characterisation of all reported molecules, all analytical methods and details about computational studies.

Acknowledgements

"Every good and perfect gift is from above, coming down from the Father of the heavenly lights, who does not change like shifting shadows". James 1:17

First of all, I would like to thank my supervisor, Dr G. Dan Pantoș, without whom I couldn't have been able to finish my Ph.D. His constant support, help and constructive critics helped me to become a better scientist every day and overcome all the problems encountered during these 4 long years. He has taught me many useful things, beyond the chemistry aspects, from how to fix HPLC instruments to more philosophical things like "the principle of throwing the shoe".

Secondly, I want to thank my lovely wife and colleague, Dora, for all her love and support during my Ph.D. Without her help, I don't think I would have finished my Ph.D.

Many thanks also go to all present and former Pantoș group members: Ana Trandafir (special thanks to Ana for her help with computational studies), Dr Mehrafarin Eilbigi Dehkordi (aka Afi), Dr Simone Limberti and Dr Giles Prentice for sharing their knowledge with me and for making the workplace enjoyable. I would also like to thank visiting PhD and undergrad students in Pantoș group from Japan and Hong Kong for great moments spent working in the lab until midnight and having fun: Yuto Kage, Ryo Nozawa and Hugo Wong. Other colleagues in the Department of Chemistry are also acknowledged: Dr Fabienne Pradaux-Caggiano, Dr Adam Sedgwick, George Williams. I am grateful to all the undergrad and Masters students I have supervised during my Ph.D. because this experience shaped my personality.

I gratefully acknowledge valuable help from the CCAF team with NMR set up (Dr John Lowe and Dr Catherine Lyall), crystallography (Dr Gabriele Kociok-Köhn) and MS set up (Dr Shaun Reeksting).

Special thanks go to my parents for all their love, support and encouragement during my Ph.D. I also thank my parents- (Decebal and Mihaela Răsădean), brothers- (Andrei Răsădean, Daniel Grigor and Daniel Puiți) and sisters-in-law (Cristina, Monica, Andreea, Ligia and Oana Răsădean) and Gabriel Grigor for their support.

I would like to thank all our friends in Bath and Bristol: Dominic, Anneliese, Dani & Denisa Bandrabur, Manu, Lidia, Naomi & Eli Bercea, Adrian Carey-Jones, Beni &

Miriam Condoroș, Cristi & Ramona Joldeș, Filip, Caleb, Cristi & Geluca Medrea, Adrian & Betty Vaida, Marius & Anca Tănasă and all the other friends in the Romanian community for their friendship and unforgettable moments. All their support is much appreciated.

I gratefully acknowledge EPSRC and the University of Bath for funding my research, and the EPSRC UK National Mass Spectrometry Facility at Swansea University for high resolution mass spec analyses.

Abbreviations

ΔG – change in free energy

1D – one dimensional

2D – two dimensional

A – acceptor

ACID – Anisotropy of the Induced Current Density

au – arbitrary units

BDF – benzo[1,2-*b*:4,5-*b*']difuran

BDT – benzo[1,2-*b*:4,5-*b*']dithiophene

CD – Circular Dichroism

COSY – correlation spectroscopy

CT – charge transfer

CT – charge-transfer

D – donor

DBU - 1,8-Diazabicyclo[5.4.0]undec-7-ene

DCC – dynamic combinatorial chemistry

DCL – dynamic combinatorial library

DMF – dimethylformamide

DMSO – dimethylsulfoxide

DN – dialkoxynaphthalene

dppf – Bis(diphenylphosphino)ferrocene

dr – diastereomeric ratio

DTT – dithiothreitol

EDC HCl – *N*-(3-Dimethylaminopropyl)-*N*′-ethylcarbodiimide hydrochloride

Equiv – equivalent

ESI – electrospray

ETFA – ethyl trifluoroacetate

HPLC – High Performance Liquid Chromatography

HRMS – high resolution mass spectrometry

HSQC – heteronuclear single quantum coherence spectroscopy

Iso-BDT – benzo[1,2-*b*:5,4-*b*']dithiophene

K – Kelvin

K_a – association constant

K_d – dimerisation constant

LC-MS – Liquid Chromatography-Mass Spectrometry

m/z – mass to charge ratio

MS/MS – Tandem Mass Spectrometry

MW – microwave

NDA – 1,4,5,8-naphthalenetetracarboxylic dianahydride

NDI – naphthalenediimide

NHS – *N*-hydroxysuccinimide

NMI – naphthalenemonoimide

NMR – nuclear magnetic resonance

NOESY – Nuclear Overhauser enhancement spectroscopy

o/n – overnight

PDI – perylenediimide

PI – pyrroloindole

r.t. – room temperature

ROESY – Rotating-Frame Overhauser spectroscopy

SDRRM – Structural Divergent Synthesis on Racemic Mixtures

$TBANO_3$ – tetrabutylammonium nitrate

TCEP – tris(2-carboxyethyl)phosphine

TFA – trifluoroacetic acid

THF – tetrahydrofuran

UV-Vis – ultraviolet-visible

VT – variable temperature

Cartoons for Chapter 2

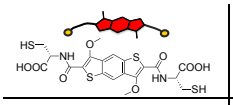
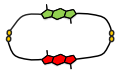
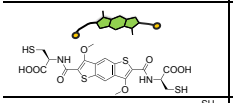
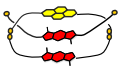
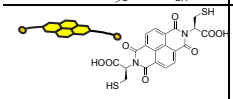
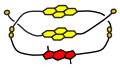
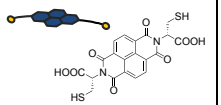
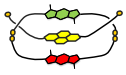
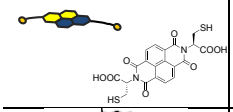
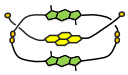
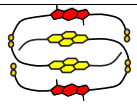
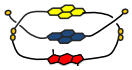
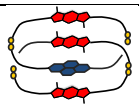
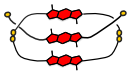
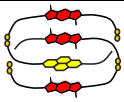
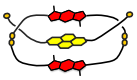
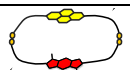
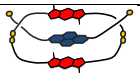
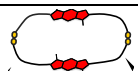
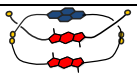
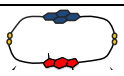
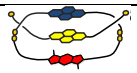
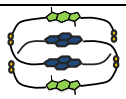
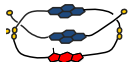
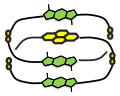
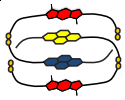
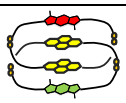
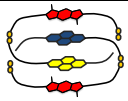
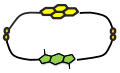
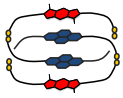
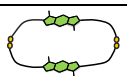
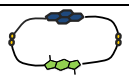
	***R,R*-2**		**Y1**
	***S,S*-2**		**E**
	***R,R*-1**		**F**
	***S,S*-1**		**F1**
	***R,S*-1**		**F2**
	**Cat I *RRRR***		**G**
	**Cat II *RSRR***		**H**
	**Cat II *RRRR***		**I**
	**X**		**J**
	**Y**		**K**
	**Z**		**L**
	**Cat I *SSSS***		**M**
	**Cat II *SRSS***		**Cat I RRSR**
	**Cat I *RRRS***		**Cat I RSRR**
	**X1**		**Cat I RSSR**
	**Y2**		**Z1**

Table of Contents

Chapter 1: Introduction

1.1. Supramolecular chemistry

Starting from the middle of the 20th century, scientists have been keen on synthesising large and interlocked molecules. These include knots, catenanes, rotaxanes *etc.*; the simplest example is [2]catenane, which consists of two single interlocked rings. The first [2]catenane (Figure 1.1) was obtained in low yield, being also the first and solely fully saturated catenane synthesised through the reductive coupling of a diester, along with other macrocycles.[1]

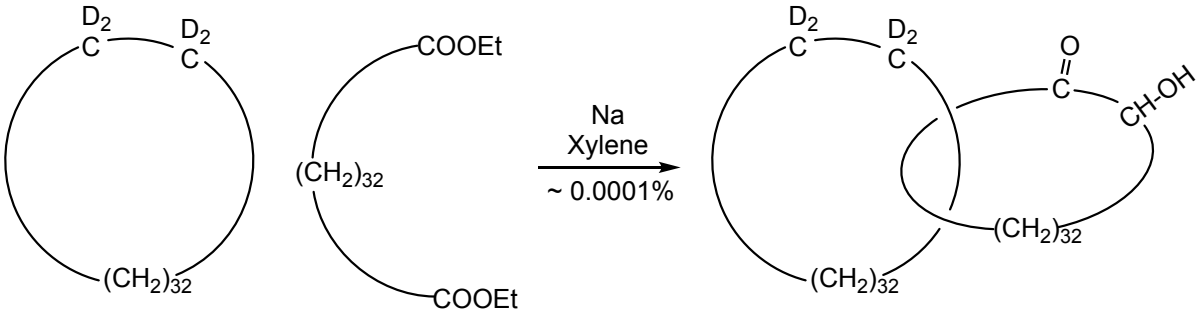


Figure 1.1. The synthetic route towards the first [2]catenane.

This was the spark that led to an extensive and challenging research to design further interlocked systems. Two different methodologies were developed over time: either using an ion as a template (cation / anion strategy) or the Dynamic Combinatorial Chemistry (DCC) approach. My work uses the DCC to making interlocked molecules, so advances in the DCC method will be detailed herein.

1.2. Dynamic Combinatorial Chemistry

One of the approaches widely employed in the synthesis of interlocked molecules is the so-called Dynamic Combinatorial Chemistry, and it involves reversible bonds. In this method, simple molecules (called building blocks) are mixed together, generating a Dynamic Combinatorial Library (DCL). The building blocks are carefully selected to favouring different interactions such as: hydrogen bonding or π-π stacking. The selection of reversible bonds (yellow dots in the cartoon representation – Figure 1.2) is made based on the experimental conditions intended to be used, such as pH and solvent. DCLs are usually stirred at room temperature until the

thermodynamic equilibrium is reached (second step in Figure 1.2). These systems are sensitive to different stimuli, such as pH variation, solvent and temperature. The distribution of a DCL can be influenced by the addition of a template (*i.e.* templation), which acts as a guest and binds to one or more components within the library, amplifying their formation (Figure 1.2).[2–5]

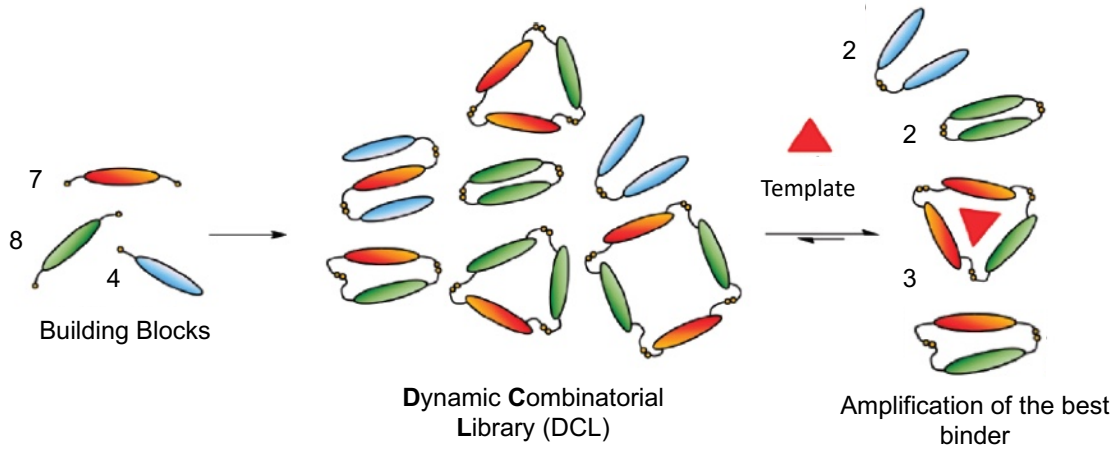


Figure 1.2. Schematic representation of DCC.[4]

Historically, DCC uses three main types of such bonds: covalent, coordinative and non-covalent (*i.e.* hydrogen bonding) bonds. Numerous types of reversible covalent bonds were employed in supramolecular chemistry, including hydrazine, imide, hemithioacetal, disulphide, *etc*. A general characteristic of these is the reversible nature of the reactions on a moderate timescale. They work under mild experimental conditions, and the exchange can be stopped at any point (before the equilibrium is reached). Most reversible covalent exchanges are limited by the pH of the environment, some of them work only under acidic conditions (for example, imine and hydrazone reactions), while others are effective at basic pH (disulphide exchange). Some of the reversible bonds most relevant to the current work are described later in this chapter. A brief account on the different types of reversible bonds involved in DCC is provided in Figure 1.3.[3]

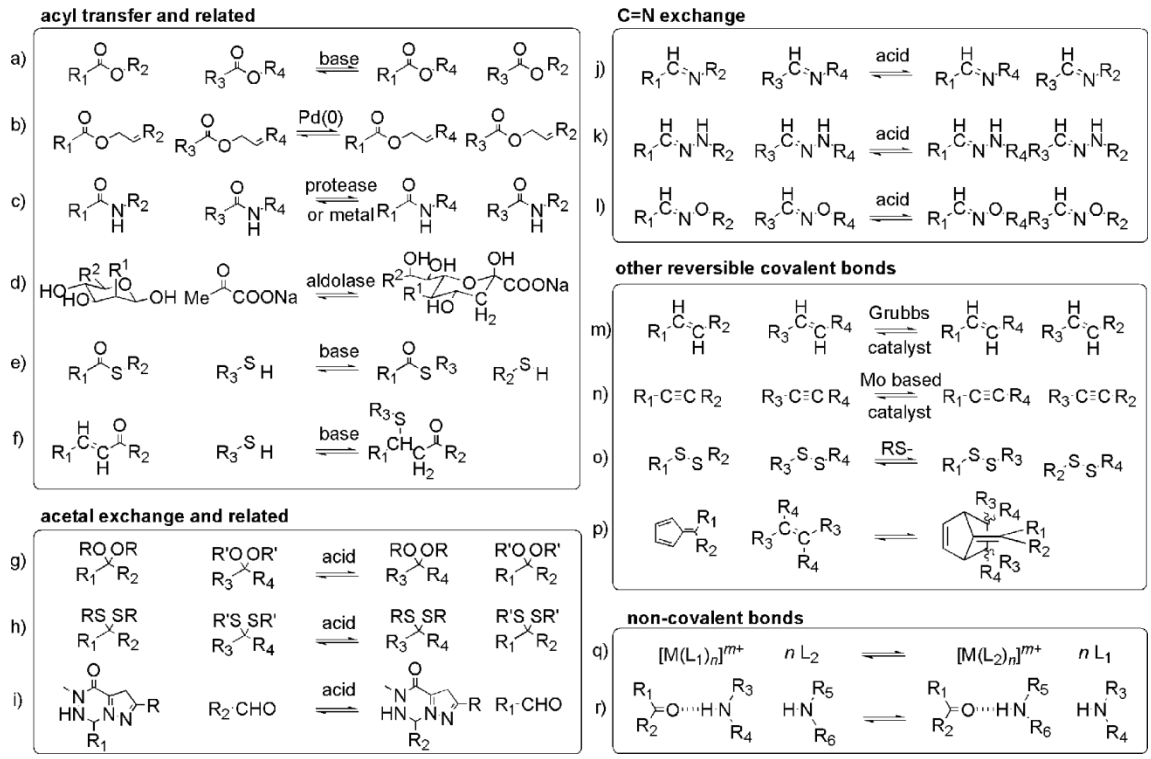
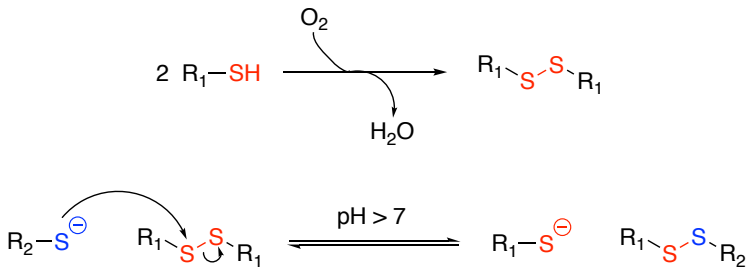


Figure 1.3. Different types of exchanging bonds: a) transesterification, b) trans-allylesterification, c) transamidation, d) aldol exchange, e) trans-thioesterification, f) Michael/retro-Michael reaction, g) acetal exchange, h) thioacetal exchange, i) pyrazolotriazone metathesis, j) transamination, k) hydrazine exchange, l) oxime exchange, m) alkene metathesis, n) alkyne metathesis, o) disulphide exchange, p) Diels-Alder / retro-Diels-Alder reactions, q) metal-ligand exchange, r) hydrogen-bond exchange.[3]

Among all of these, the disulphide bond stands out as it resembles the linkages within biological systems (*e.g.,* S-S linkage in proteins).[6–10] The exchange mechanism is reversible, pH-dependant (aqueous media at pH around 8), and it involves deprotonation of a thiol group. The thiolate anion attacks an existing disulphide bond (formed by the oxidation of two thiol moieties); when all thiolate ions are consumed, the exchange stops (Scheme 1.1).



Scheme 1.1. Disulphide bond formation / exchange mechanism.

1.3. Dynamic combinatorial chemistry using disulphide bonds

The next section describes the advancement in DCC through disulphide exchange. This area of chemistry was extensively used by the Sanders' group to generate interlocked structures. The driving forces behind the formation of such molecules are mainly π-π interactions and the hydrophobic effect,[2,4] which are briefly introduced below.

1.4. π-π interactions

Non-covalent chemical interactions are ubiquitous in nature and the human body (such as hydrogen bonding networks in-between DNA-bases), playing a key role in our daily life. There is a plethora of types of chemical interactions, but hydrogen bonding and π-π interactions are some of the most common ones.[5]

Aromatic interactions were loosely defined until Hunter and Sanders proposed an electrostatic model that explains this interaction.[11] The approach describes the way by which aromatic molecules stack together, so the optimum interaction is achieved. Under this assumption, an aromatic molecule is understood as having a partial positive charge surrounded by two negative partial ones. There are several types of stacking interactions between aromatic molecules, depending on their electronic properties (electron-rich or electron-deficient). The idea is depicted in Figure 1.4, where cores of opposite electronic properties are shown. The two cores correspond to the well-studied electron-deficient naphthalenediimide (NDI) and electron-rich dialkoxynaphthalene (DN) molecules. Subsequently, Wheeler and Houk countered the Hunter / Sanders' model using computational approaches, suggesting that the stacking is influenced by electrostatic interactions through space due to polarising effect induced by aromatic substituents.[12]

The so-called face-centred stacking occurs in systems having both a π-donor and a π-acceptor moiety. The stacking is facilitated by the electrostatic attraction between the partial positive and negative charges of the acceptor and donor, respectively.

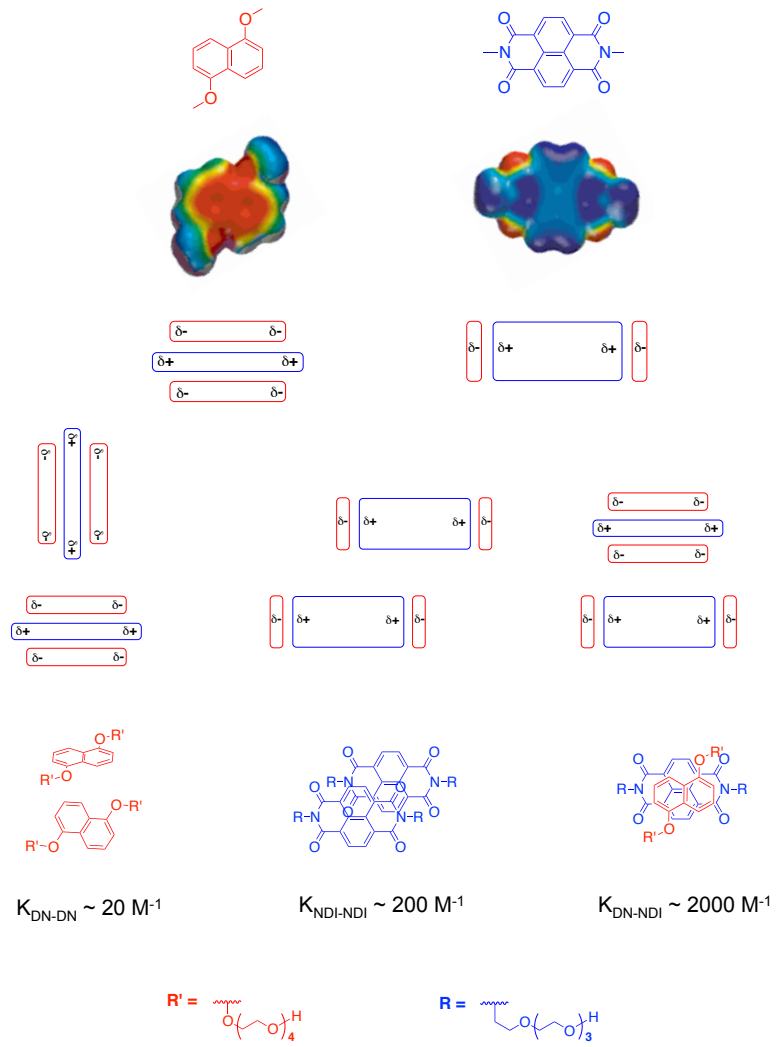


Figure 1.4. The electrostatic model of aromatic interactions. Top: calculated electrostatic surface potential for 1,5-dimethoxynaphthalene and dimethyl NDI (red denotes areas of relatively high electron density; blue stands for electron deficient areas). Bottom: possible stacking arrangements observed in crystal structures and the corresponding association constants in water as reported by Iverson *et al*.[13]

In the case of two π-rich aromatic molecules, the interaction is highly unfavourable due to the repulsion of the negative charges. Another type of π-π interactions is represented by the off-centre parallel stacking as a feature of two acceptors coupled together. The π-acceptor molecules are susceptible to edge-to-face interactions recognised by a T-shaped arrangement. Among these, the face-centred stacking favours most the interaction between the clouds of π electrons of the components. A comprehensive review on this topic relatively recent reported by

Iverson *et al.* describes these as “aromatic donor-acceptor interactions” rather than “π-π stacking”.[14]

1.5. Hydrophobic effect

For understanding the hydrophobic effect, one needs to visualize the impact that solutes have on the structure of water. In all cases of solvation, a cavity is formed and the solvent molecules need to reorganise, so to keep the original system as intact as possible. In the case of the hydrophobic effect, the water molecules arrange in a basal orientation, in which the proximity of electrons, oxygens atom, and lone pairs are maximised. As for the thermochemical aspect of the hydrophobic effect, the assemblies are entropically driven because water molecules are released from an ordered arrangement towards the hydrophobic groups.[15]

After the introduction of the main concepts, the discussion focuses on DCC using disulphide exchange. The building blocks used in DCLs are cysteine-modified naphthalenediimides (NDI) and dialkoxynaphthalenes (DN) as π-acceptors and π-donors, respectively (Figure 1.5).

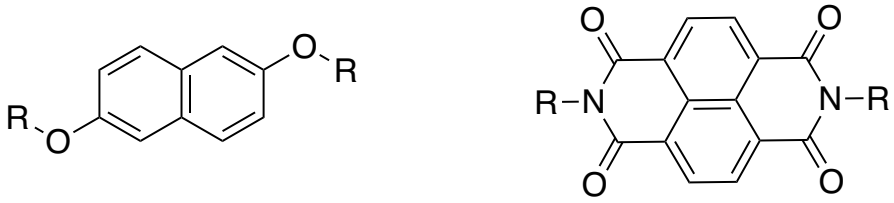


Figure 1.5. The general representation of a dialkoxynaphthalene (π-donor) and a naphthalenediimide (π-acceptor).

DCC using disulphide exchange has been efficient in synthesising numerous interlocked assemblies and novel compounds from simple [2] and [3]catenanes to elaborated structures as trefoil knots. One of the simplest [2]catenanes obtained *via* DCC is solely based on DN and NDI scaffolds. However, a [2]catenane can adopt more complex structures, depending on the complexity of the building blocks. The driving force behind catenanes’ formation is the hydrophobic effect rather than the donor (*D*) – acceptor (*A*) interaction. Sanders’ group has come up with separate pathways towards to formation of different types of [2]catenanes (Figure 1.6). The

size of the cavity of initial dimers can be tailored by changing the length of the side chain of acceptor and donor building blocks as well as their connectivities (*i.e.* 1,5-DN *vs.* 2,6-DN). This leads to the formation of [2]catenanes with different stacking arrangement: *DADD*, *DADA*, *DAAD*, *AADA*.[16] The discovery of different types of catenanes was a breakthrough, particularly because researchers have considered, up to that point, that the only favourable interaction would be the *DADA* stacking, which is the most favourable. This strengthens again the important role played by the hydrophobic effect in disulphide exchange-based DCC. However, at that stage, the synthesis of all-acceptor or all-donor catenanes from libraries containing donor or acceptor building blocks was considered, unsurprisingly, a challenge.

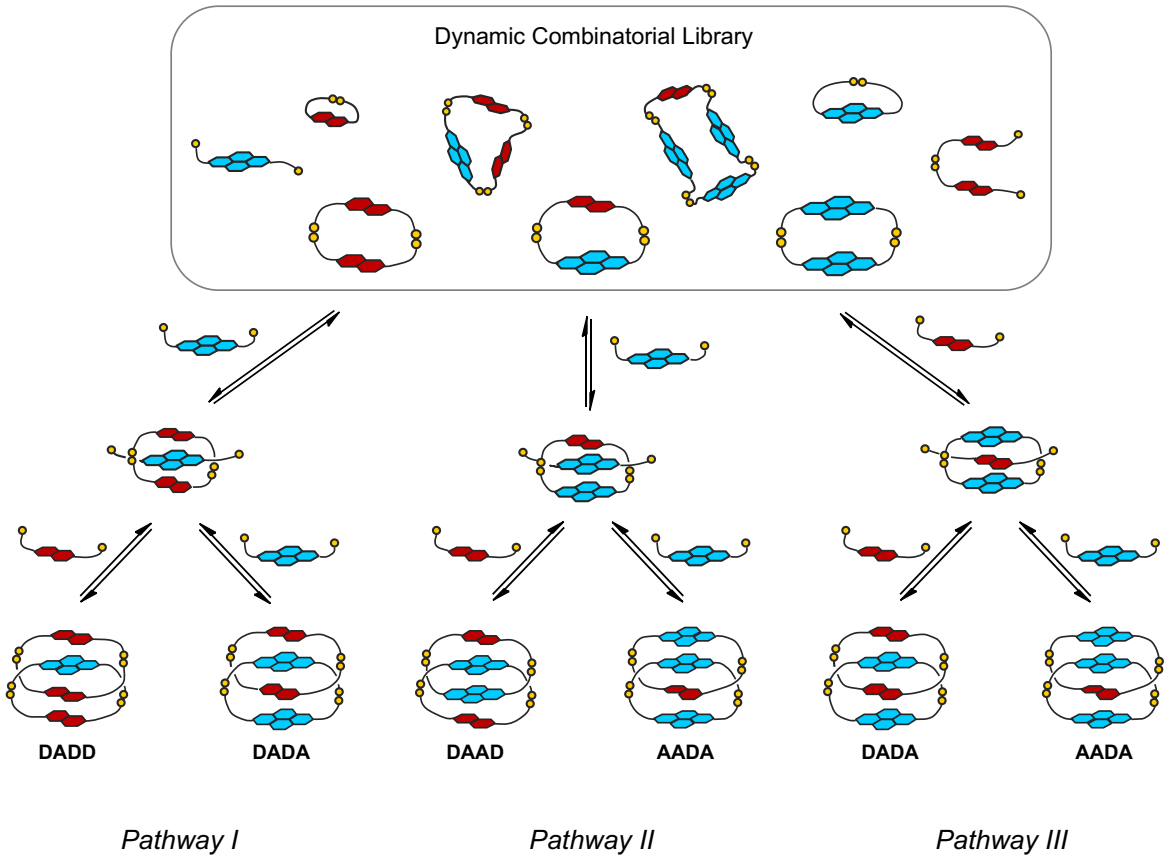


Figure 1.6. The proposed mechanism for the formation of different [2]catenanes using the DCC method. The yellow dots represent sulphur atoms.[16]

Over time, the complexity of structures obtained by this approach was enhanced. New NDI-based building blocks were explored, such as one with two NDI cores linked by aliphatic scaffolds, which contained up to nine carbon atoms between the two NDI cores. The length (and nature) of the linker was varied to test the effect of rigidity of the system.[17] This led to the formation of the first [3]catenane and an all-acceptor [2]catenane based on disulphide exchange. The whole process is described in Figure 1.7. The all-acceptor [2]catenane was obtained from a library containing only an acceptor building block with at least eight carbon atoms between NDI cores. The driving force was attributed to the hydrophobic effect. Based on the ^{1}H NMR spectrum, the linker was "buried" in the cavity, whereas the hydrophilic cysteine part pointed towards cavity's exterior. Another work from Pantoș' group showed the influence of nitrogen atoms in the linker.[18] The linkers used in the study were spermidine and *N,N'*-bis(2-aminoethyl)-1,3-propanediamine. In both cases, the [2]catenanes formed in high yield, with the spermidine-based [2]catenanes having eight isomers because of the intrinsic planar chirality of the DN dimer and the arrangement of NDI core inside the DN dimer.

NDI-based acceptor building blocks with less than seven carbon atoms have exhibited an odd-even effect.[17] Therefore, the building blocks with two, four and six carbons in the linker have led to [3]catenanes formed in good yields. Based on kinetic studies, the acceptor dimer formation was the first step followed by further addition of two donor dimers. However, for linkers containing an odd number of carbon atoms, a [3]catenane was not formed in significant concentrations (Figure 1.7). When the linker contains over six carbon atoms, the odd-even effect is less expressed, thus *DADA*-type [2]catenanes are formed.

A DCL containing only the NDI building block with six carbons in the linker was set up. The DCL distribution showed the formation of a dimer and a trimer as well as four unexpected tetramers (tetramers **1.1** – **1.4).** Based on its retention time, the tetramer **1.4** was identified as a macrocycle. The rest of the species were identified as interlocked molecules as following: a trefoil knot **1.1**, a Figure Eight Knot **1.2** and a Solomon link **1.3** (Figure 1.8).[19]

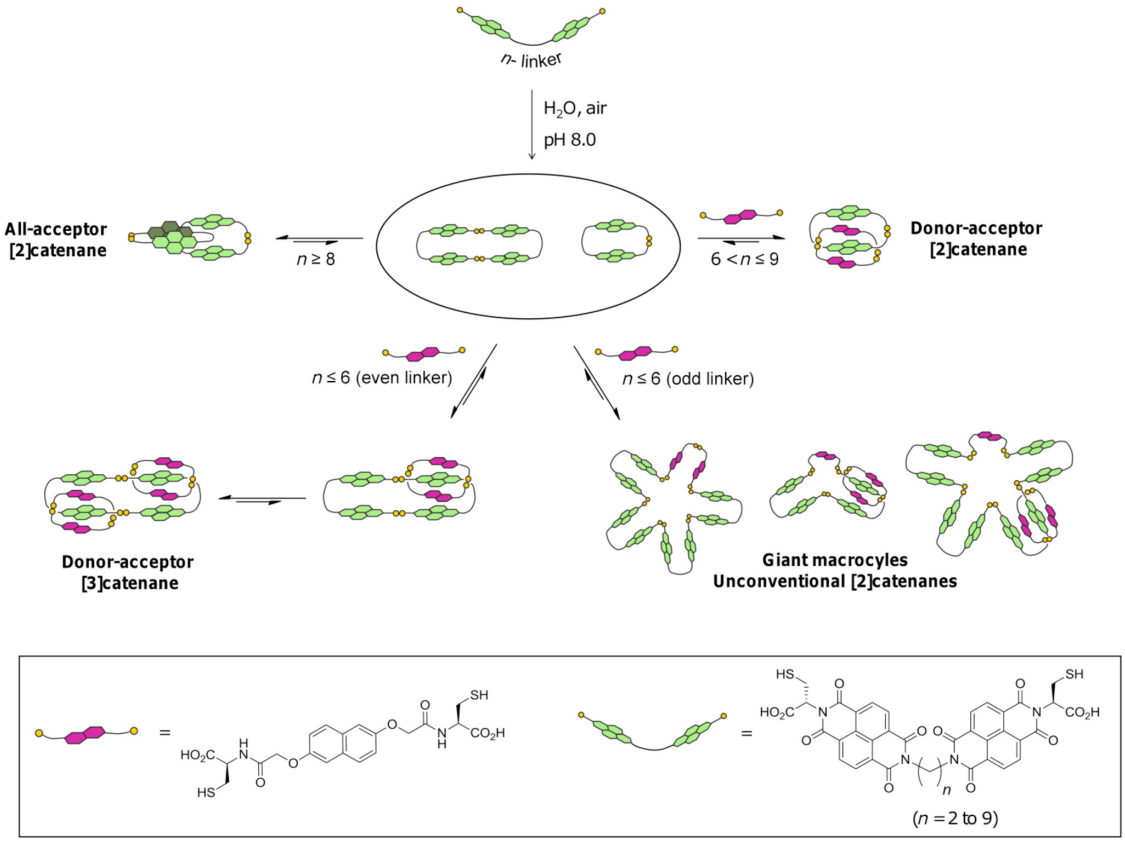


Figure 1.7. The formation of different [2], [3]catenanes and macrocycles using DCC. The yellow dots represent sulphur atoms.[17]

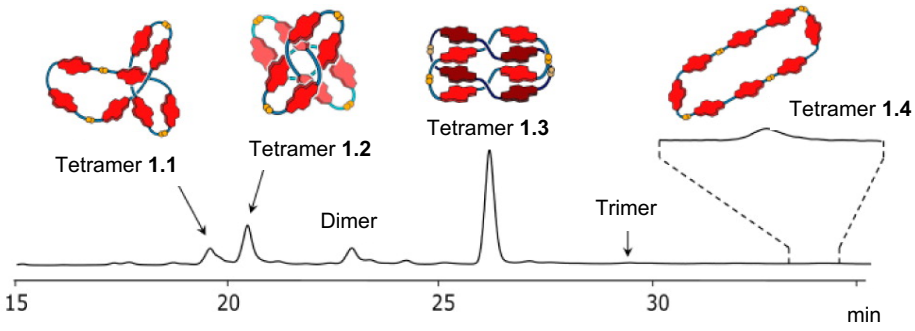


Figure 1.8. Different topologies (links and knots) formed within a library containing linkers with six carbon atoms. The yellow dots represent sulphur atoms.[19]

Expanding on this theme and adding a carboxylic acid group to a linker with two carbon atoms, a trefoil knot (Figure 1.9) was quantitatively obtained.[20] The driving force of the reaction was also the hydrophobic effect. The polarity of the reaction mixture required to increase the hydrophobic effect was reached by adding salts in different concentrations. When tetrabutylammonium nitrate was used, which can solvate the NDI core, the dimer was formed at the expense of the knot. The same

behaviour was observed for the aforementioned catenanes. The information offered by the kinetics of synthetically formed knots paves the way towards understanding the more complex protein-based knots from the biological world.

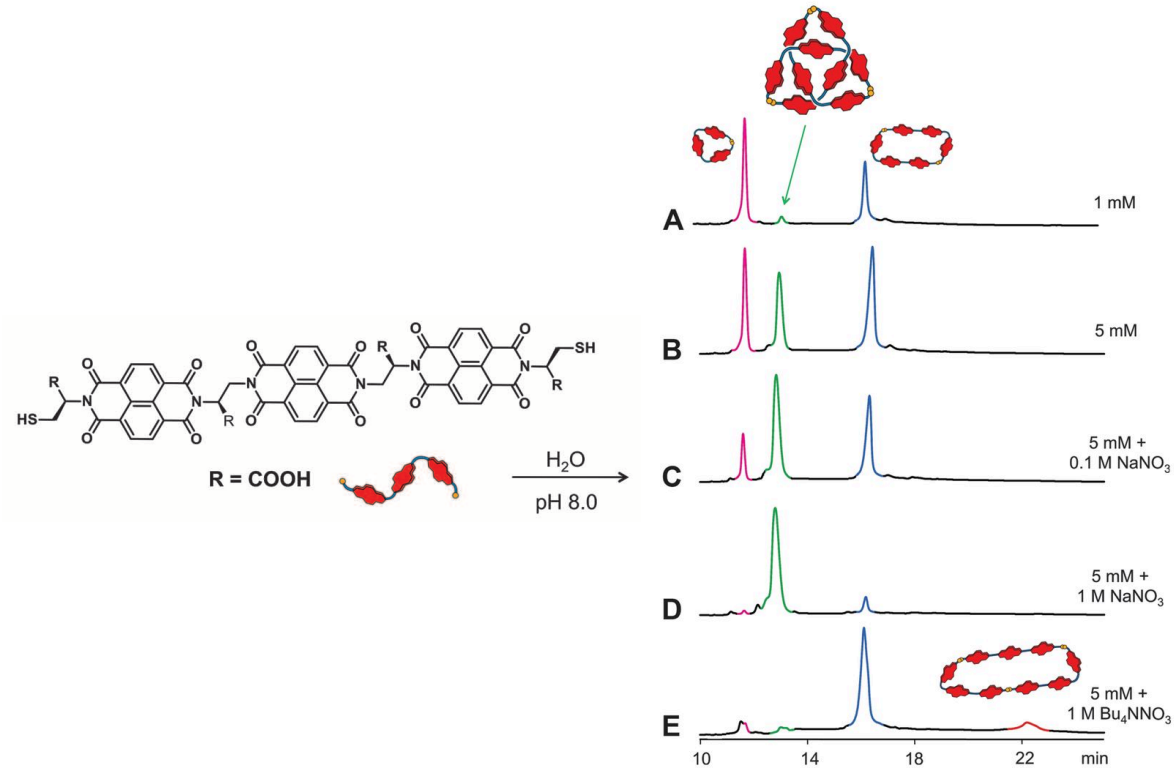


Figure 1.9. a) HPLC analyses of NDI trimer libraries at concentration of 1 mM, b) 5 mM and in the presence of different salts: c) 0.1 M $NaNO_3$, d) 1 M $NaNO_3$ and e) 1 M tetrabutylammonium nitrate. The yellow dots represent sulphur atoms.[20]

Recently,[21,22] Sanders *et al.* have used the disulphide exchange to yield molecular cages in water based on benzenedithiol and benzenetrithiols. When the building blocks were mixed without template, just trimers, tetramers and small cages derived from benzenetrithiol were formed. The thermodynamic stability and entropically favoured formation of benzenetrithiol-based dimers have led to a simple (trivial) library distribution (Figure 1.10). On the other hand, when different templates (*i.e.* spermine, spermidine*, etc.*) were added, the library distribution changed. Several new cages were formed in water; some of them of large sizes, containing up to nine components were identified in the system. Also, the cage formation was dependent on the template concentration, thus, the lower the concentration, the lower the yield of the cage was. However, due to the system complexity and very close retention times of forming species, the cages could not be separated.[21]

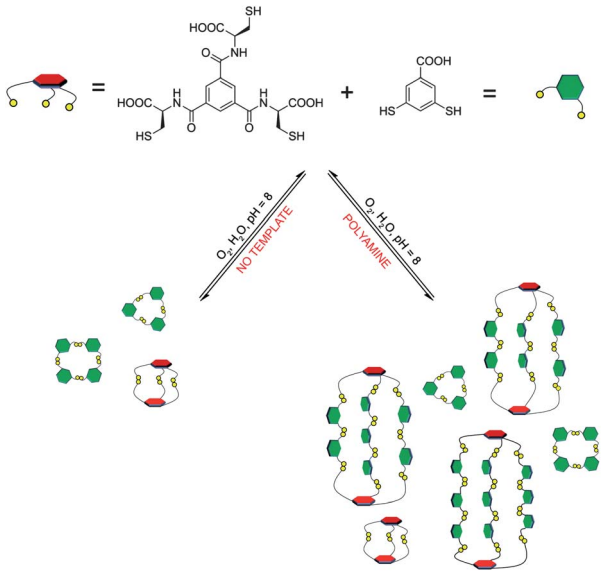


Figure 1.10. Library composition identified within the DCL of the trithiol with no template as well as a templated DCL with polyamine leading to cages formation.[21]

Otto *et al.* have discovered a self-replicating molecule forming peptide β-sheets fibres based on benzenedithiol. Different products have formed under various conditions: the library was shook (hexamer) or stirred (heptamer). A hypothesis for this behaviour was offered based on different energies facilitated by these two processes (stirring generates higher energy than shaking; Figure 1.11).[23,24]

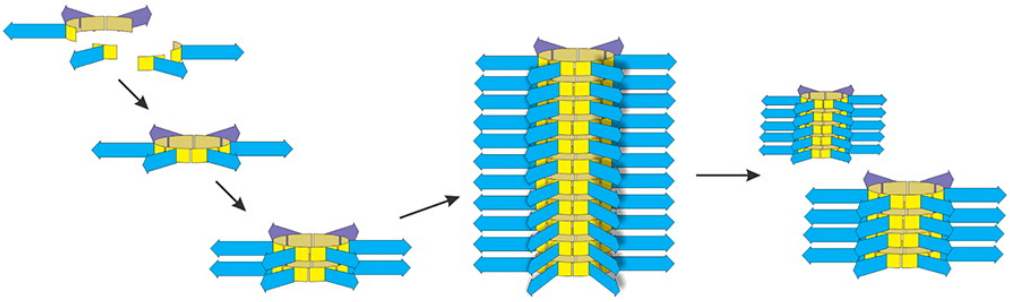


Figure 1.11. β-sheets forming entities composed of self-assembly peptides that shape into fibrils.[25]

More recently, Otto's group have reported a multi-component system based on two benzenedithiol peptides, phenylalanine and serine (Figure 1.12).[26] Initially, the phenylalanine-based peptide was consumed forming hexamers, while the serine-functionalised peptide had preferentially generated trimers and tetramers mainly due to serine's hydrophilic nature. The system was investigated ten days later

and new serine-peptide hexamers were formed, which subsequently evolved into a second generation of hexamers. This phenomenon was ascribed to a cross-catalysis process within the different components of the library.

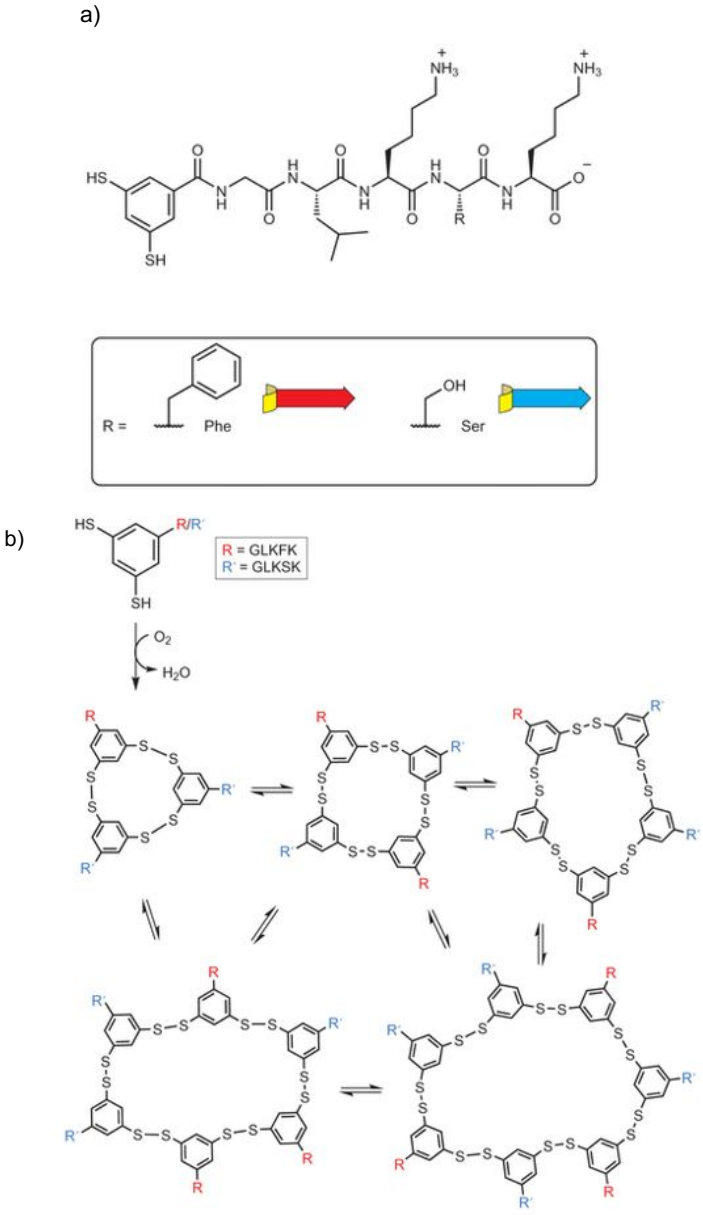


Figure 1.12. a) The structures of building blocks involved in the formation of self-assembly peptides, b) partial distribution of the DCL formed by benzenedithiol peptides, phenylalanine and serine building blocks.[26]

Besides synthesising interesting interlocked architectures, DCC is also useful to finding receptors for biological molecules as well as in anticancer therapies. Thus, DCC has allowed the development of highly specialised molecules towards recognition of a particular entity. For example, a research group from North Carolina

have discovered a molecular receptor for dimethyl arginine.[27] Such receptors are of particular interest as irregularities in the methylation process of arginine trigger leukaemia or breast cancer. The building blocks (annotated as **1.5**, **1.6** and **1.7** – Figure 1.13) were designed in such way to tailor the distribution of the library. The species obtained so – (**1.5**)$_2$(**1.6**) and (**1.5**)$_2$(**1.7**) – exhibited selectivity towards dimethylated arginine. Therefore, (**1.5**)$_2$(**1.7**) species has been amplified in the presence of dimethylated arginine, whereas the cavity binding site has remained unchanged in the presence of arginine or the monomethylated derivative. This can be understood based on either the difference in size between various stages of methylated molecules, or the available surface area for interactions.

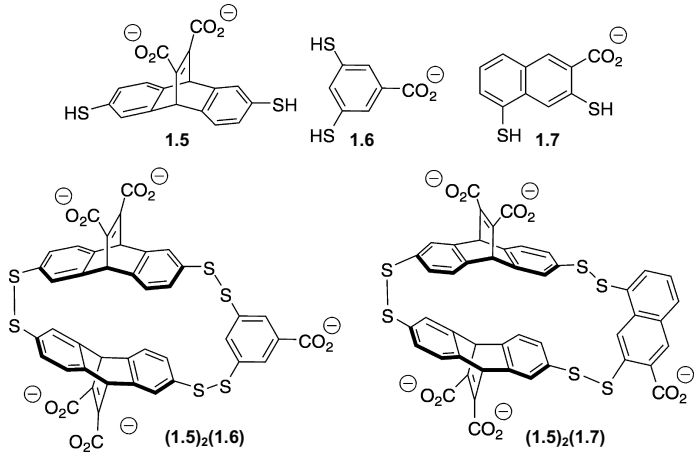


Figure 1.13. Structures of building blocks employed in DCL for arginine receptor (**1.5**, **1.6** and **1.7**), and two of the components found in the library.[27]

Disulphide exchange has led to the synthesis of a large part of interlocked molecules reported so far in the literature such as rotaxane,[28] cage,[21] molecular arm[29] *etc*. The formation of [2] and [3]rotaxanes through disulphide exchange has been reported by Takata's group.[28] The synthetic pathway assumed firstly the axle formation with 3,5-di-*tert*-butylphenyl as stoppers, and a disulphide moiety with two secondary ammonium salts in the centre of the rotaxane. The exchange was favoured by the addition of a small thiol-based molecules, allowing for the "wheel" (dibenzo-[24]crown-8) to bind to the ammonium group, leading to the aforementioned rotaxanes (Figure 1.14).

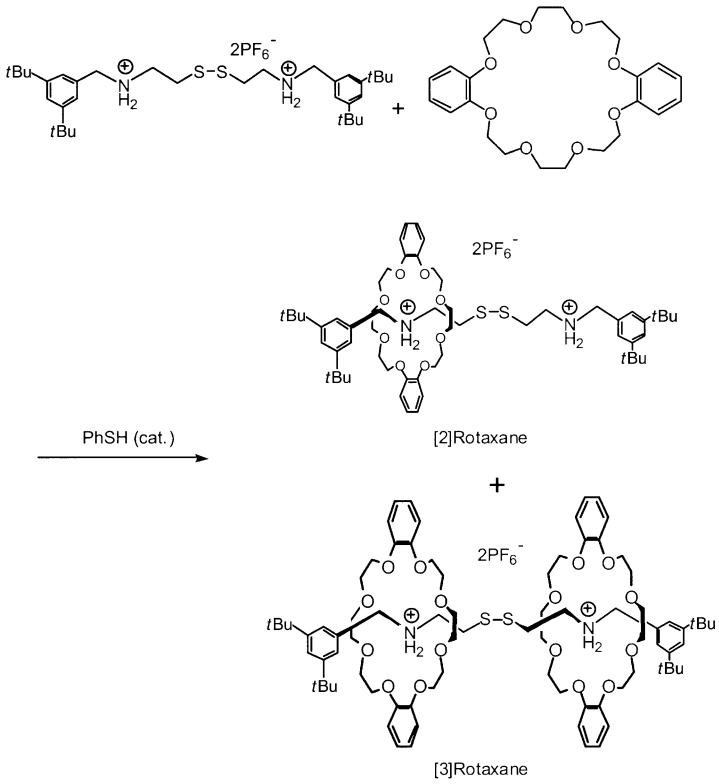


Figure 1.14. The synthetic scheme for the synthesis of [2] and [3]rotaxane designed by Takata's group.[28]

Foldamers are another representative class of molecular architectures obtained through disulphide exchange. Their formation and stability have been explored by Ivan Huc *et al.*, who designed two oligoamide and thiol-based foldamers with C_2 symmetry.[30] One foldamer forms *via* an intramolecular disulphide bond, whereas the other one is a difoldamer linked by an intermolecular disulphide bond. The latter one shows a strong preference towards one conformation (*MM* or *PP*) due to the arrangement of the thiopropoxy scaffold, although the helices do not communicate (Figure 1.15). The intermolecularly linked foldamer has exhibited an almost double half-life relative to racemisation in comparison with the intramolecularly formed helices.

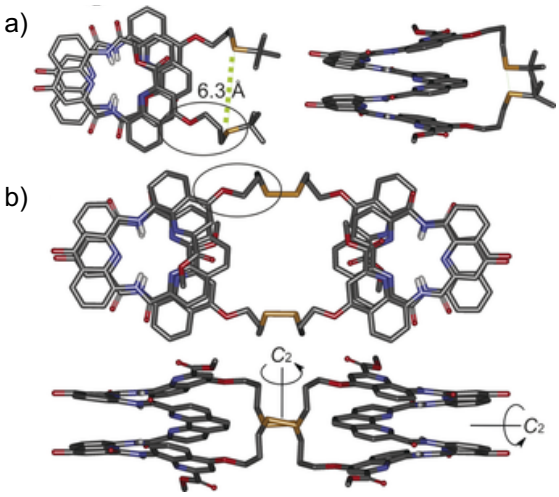
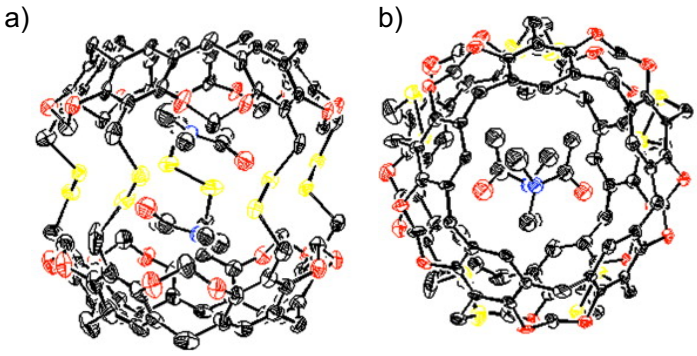


Figure 1.15. The crystal structure of a) foldamers' precursor and b) the foldamers linked intermolecularly *via* hydrogen bonding.[30]

Carceplexes are another class of topologically complex molecules, consisting of two cavitands linked by two or more covalent bonds. Sherman *et al.* have obtained [4] and [5]carceplexes connected through four and five disulphide bonds, respectively.[31] These molecules are also able to accommodate small guests, such as thioethers or ketones. The cavity of the [4]carceplex accommodates only one guest, whereas the [5]carceplex has a larger cavity, thus accommodating two guests (Figure 1.16).



1.8

Figure 1.16. The crystal structure of a) [5]carceplex and b) [5]carceplex embedding two molecules of dimethylacetamide.[31]

Disulphide exchange has also led to other classes of compounds: replicators[23,32] developed by Otto's group, or various receptors studied by Waters'

group[27] or Balasubramian's group;[33] but these go beyond the aim of the present review.

1.6. Other reversible bonds

1.6.1. Imine exchange

DCC also uses other reversible bonds besides the disulphide exchange as described in section 1.2. Imine and hydrazone exchanges are some of the most known reversible reactions for synthesising interlocked molecules through DCC. In contrast with disulphide exchange, they are rather effective under acidic pH, and inert in basic conditions. Such exchanges have led to the formation of Borromean rings,[34] Solomon links,[35,36] rotaxanes,[37,38] cages[39–45] *etc.*; the most relevant examples are highlighted below. These include the first ever synthesised molecular Borromean ring starting from a mixture of six equivalents of a diamine, ketone and a Zn^{II} (or Cu^{II}) source, as reported by Stoddart *et al.* (Figure 1.17).[34] The large number of π-π interactions and dative bonds (*i.e.* 12 and 30, respectively) is the driving forces behind its formation. The crystal structure shows that each ring adopts a chair-like conformation. In the absence of a metal ion (*i.e.* the template), a mixture of macrocyclic and polymeric products has been obtained, highlighting the key role played by the metal ion in the assembling process. By simply mixing two different metal ion sources (Cu^{II} and Zn^{II} ions), a Solomon link rather than a Borromean ring was formed; its formation was thought to be kinetically driven.[35] The mass spectrometry analysis of solution from a single crystal has confirmed the formation of the Borromean species. Its dissolution confirmed the presence of borromate species, indicating an equilibrium between Solomon link and Borromean ring *via* imine exchange.

The imine exchange has also led to the formation of rotaxanes, as reported by Stoddart's group. For the design of the axle, a secondary alkyl ammonium scaffold as a recognition site for crown ether was used, and 3,5-di-*tert*butyl or dimethoxyphenyl groups as stoppers. The ring has closed around the axle *via* the imine exchange between a dialdehyde and diamine, yielding the rotaxanes.[37,38,46] The same synthetic procedure has allowed for the syntheses of [3], [5] and [7]rotaxanes. The [5]rotaxane has included four NDI moieties arranged in a cofacial manner (acting as electron transfer units). The versatility of this approach has also enabled the formation of a

[c2]daisy chain, in which two molecules possessing ammonium and crown ether moieties interact.[47]

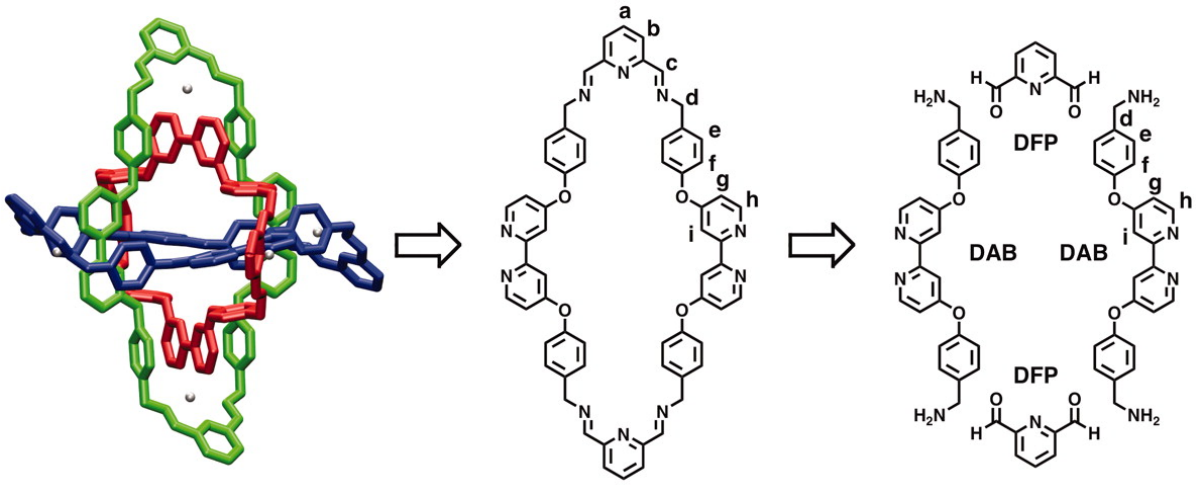


Figure 1.17. Retrosynthesis analysis of the Borromean rings.[34]

An important part of the review should be attributed to cages formed *via* other than disulphide exchange reversible processes. Cooper's group obtained organic cages, starting from an aromatic trialdehyde and a diamine.[45,48] A short diamine-based linker (1,2-ethylenediamine) and 1,3,5-benzentricarboxaldehyde have yielded a non-porous cage. When the longer 1,2-propylenediamine was used, 1D channels were formed, indicating the importance of the linker length. In a different style, a cyclic linker such as (*R,R*)-1,2-diaminocyclohexane has led to porous cages[45] that assemble with iodine or osmium tetraoxide.[49] These molecules have found applications in separating enantiomers as well as inert gases (Xe) in low concentrations from air.[50] The aforementioned trialdehyde with 1,2-ethylenediamine forms, in the presence of TFA, a triply interlocked cage; this has been also obtained when mixing the non-interlocked cage with *p*-xylene in dichloromethane. The driving force behind the formation of these molecules was π-π staking, as illustrated in Figure 1.18.[44]

The chemistry of cages has received increased attention due to their interesting arrangements as well as the possibility of further functionalisation. Such cages have been shown to stabilise reactive molecules, promote the encapsulation of gases or small molecules, favour the molecular separation of gases and porous liquids, and also act as catalysts. There is a plethora of cages-based molecular architectures reported in the literature; we highlight just a few other representative examples.

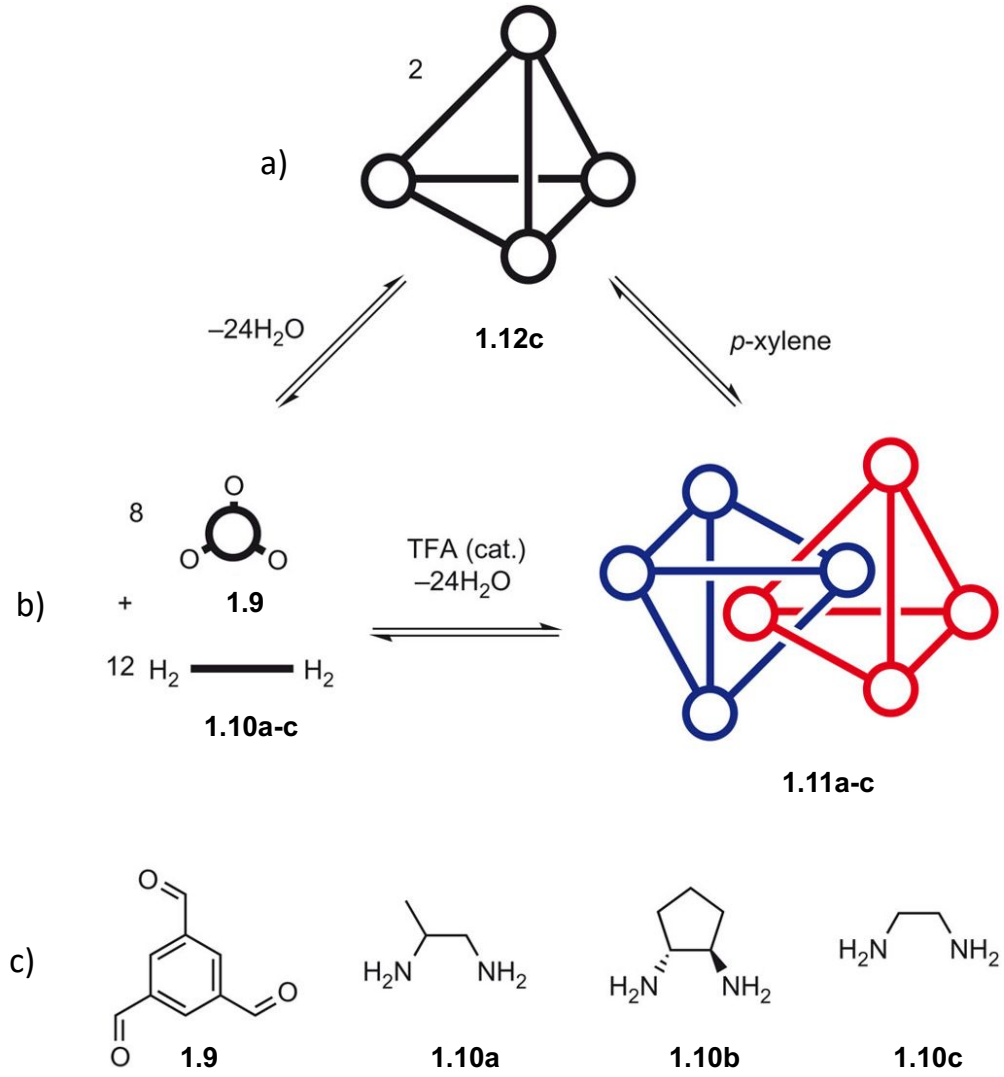


Figure 1.18. a) Schematic representation of a non-interlocked cage, b) Schematic representations of starting materials and the triply interlocked [2]catenane, along with reaction conditions and c) chemical structure of starting materials for the synthesis of triply interlocked [2]catenane.[44]

The first example comes from Liu and Warmuth, which have developed a solvent-dependent system, in which a wide range of cages is formed by varying the solvent from THF to CH_2Cl_2 and $CHCl_3$.[51] The synthetic route involves a condensation step between 1,2-ethylenediamine and a cavitand bearing four aldehyde groups. In the presence of THF, the cage formed is a result of mixing four cavitands and eight amines that leads to a distorted tetrahedron. Changing the solvent to CH_2Cl_2, a [16+32] condensation occurs, yielding a square antiprismatic molecule; a [6+12] condensation in $CHCl_3$ gives an octahedron. Despite CH_2Cl_2 and $CHCl_3$ having similar

properties, the outcome is distinct; the authors have attributed this behaviour to different solvability of the ligands under these conditions.

Cages are prospective tools for performing reactions inside enclosed spaces, because the reactive compounds can be stabilised by trapping them within the cavity. Thus, the condensation of 4-tert-butyl-2,6-diformylphenol with 1,3,5-triaminocyclohexane in a [12+8] fashion has allowed the synthesis of an octahedral cage with 3 nm diameter. Providing that entropy has played a decisive factor in the formation of this cage, the yield was nearly quantitative (Figure 1.19).[52]

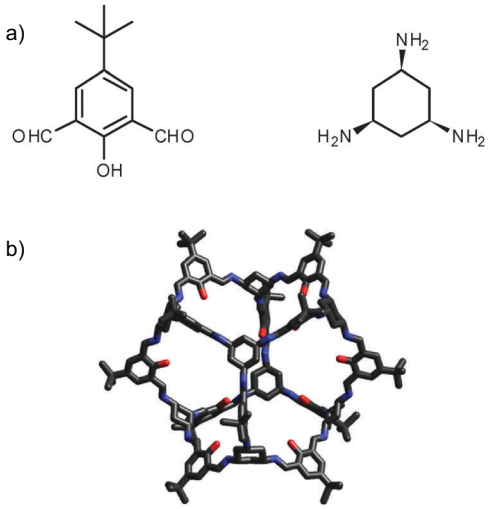


Figure 1.19. a) Starting materials for the synthesis of the octahedral cage *via* imine exchange, b) structure of octahedral cage.[52]

The combination of imine exchange with metal complexation is another approach towards the development of cage-like structures. The Nitschke's group have applied this method to synthesise metal-cages **1.13**, which upon threading with DN-based crown ethers **1.14** form [n]catenanes [(**1.13**)(**1.14**)$_n$ n = 2 – 7; Figure 1.20}.[40] The policatenated species have adopted a non-cooperative binding, assuming that the cage is large enough to thread all the crown ethers. The replacement of NDI scaffolds with nickel-porphyrin and 2-formylpyridine has enabled the accommodation of C_{60} and C_{70} within the tetrahedral cage. The system has shown an enhanced flexibility as it could be further functionalised by the addition of Cu^{I} salts, leading to hetero-metallic cages.[53] For determining the most thermodynamically favoured pathway in the presence of different templates, NDI-porphyrin-based molecules were

added to the library. The introduction of a C_{60} molecule has pushed the equilibrium towards the formation of pure NDI and templated porphyrin cages. In contrast to this behaviour, the addition of DN-based crown ether as template has generated a mixture of NDI-porphyrins cages. This experiment emphasised the ability of porphyrin-forming tetrahedrons to accommodate fullerenes, the size of the cage plays an important role in accommodating either C_{60} or C_{70}.[41]

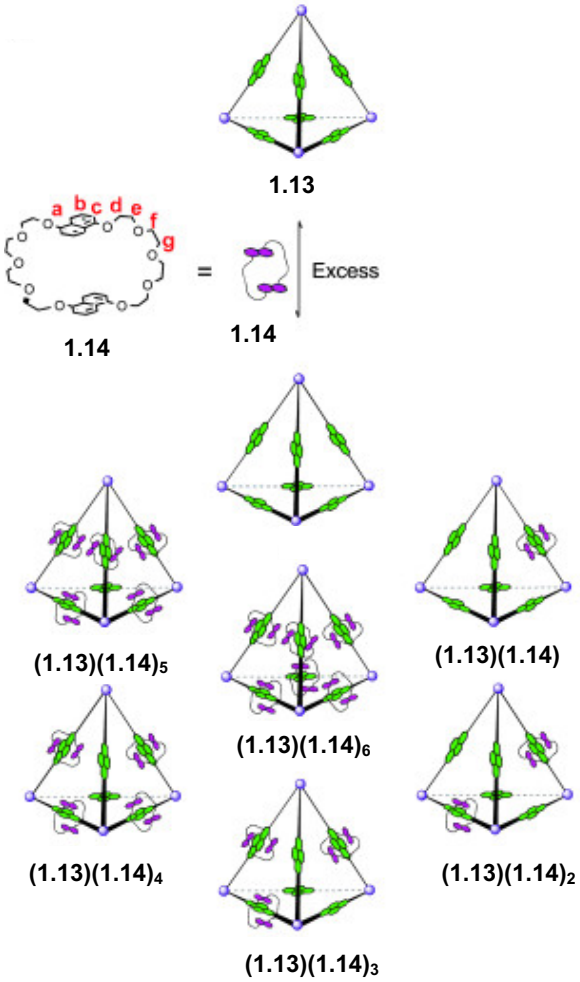


Figure 1.20. The formation of the polycatenated structures starting from the cage (**1.13**) and DN dimer (**1.14**).[40]

The imine exchange has also facilitated the formation of other synthetically challenging architectures such as knots and catenated species. For example, a [3]catenane was obtained *via* the preorganisation of di- and tetrapyridines around a Fe^{II} cation, and subsequent reaction with a diamine; the demetalation releasing the fully organic [3]catenane.[54] Another representative molecule obtained under a similar procedure, but using only one type of ligand, is the pentafoil knot reported by Leigh's

group.[55] The synthetic pathway involved a stepwise approach, initially generating the "open" knot followed by its reaction with a diamine to yield the close metalated knot (molecule confirmed by crystal structure). The cavity of the knot has accommodated a chloride ion, which is strongly bound through ten CH···Cl interactions.[55] Introduction of a simple pair of ligands and a metal cation has led to trefoil knots as well as a [2]catenane and a Solomon link.[56]

1.6.2. Hydrogen bond exchange

Besides the disulphide and imine chemistries outlined so far, other reversible reactions leading to topologically complex molecules are highlighted herein. To start with, we introduce a dynamic system composed of hydrogen bonding networks developed in Sanders' group a decade ago. It was shown that amino acids-functionalised NDIs **1.15** tend to self-assemble into helical nanotubes, a process driven by carboxylic acid dimerization and non-classical C-H···O hydrogen bonding networks. The addition of methanol triggered the disruption of the hydrogen bonding system into monomers, accounting for the hydrogen donor and acceptor nature of this solvent. On the other side, the use of C_{60} as template has led to longer nanotubes than previously identified. In the case of C_{70}, a larger analogue of C_{60}, a different mechanism has driven the assembly: the molecules re-organised into a capsule instead of folding into a supramolecular nanotube (Figure 1.21).[57–59]

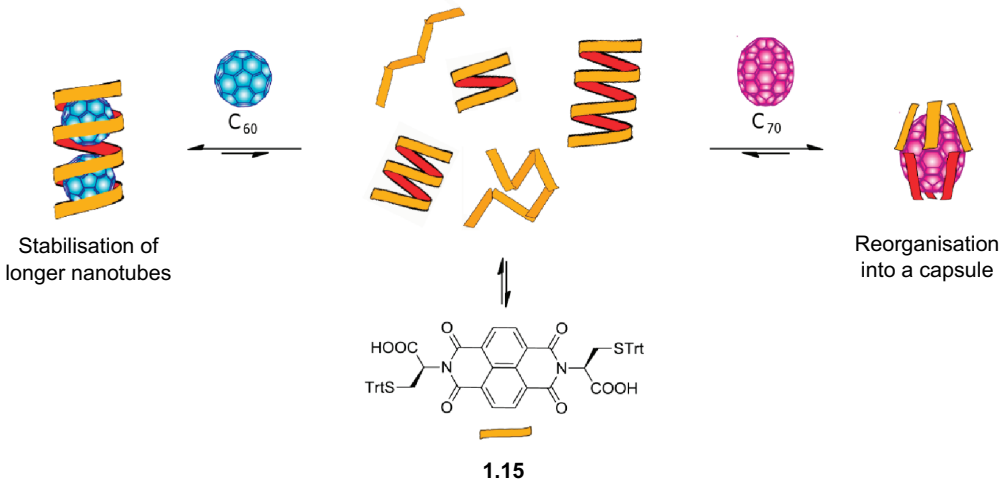


Figure 1.21. The re-organisation equilibrium between nanotubes and capsule of **1.15**.[4]

1.6.3. Metathesis and Boronic esters exchange

Returning to cages, besides imine exchange, other reversible reactions were employed in synthesising cages. A representative example is the formation of cages *via* metathesis (either alkene or alkyne) in a DCC-like fashion. This type of metathesis assumes a reversible process, unlike other metathesis strategies, for obtaining interlocked molecules. Using reversible metathesis chemistry and computational modelling, the compound shown in Figure 1.22 was synthesised and used for binding fullerenes. The UV-Vis studies suggested association constants of 1.4 x 10^5 M^{-1} and 1.5 x 10^8 M^{-1}, for binding C_{60} and C_{70}, respectively, in a 1:1 binding fashion. The difference in the strength of interaction was linked to a better match of the size of the cage with C_{70} compared to C_{60}.[60]

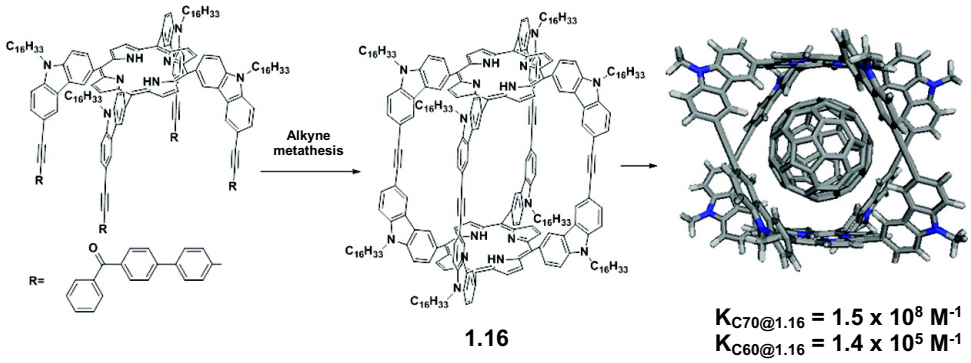


Figure 1.22. The alkyne metathesis formation of the cage, its computational model with C_{70}, and the association constants with C_{60} and C_{70}.[60]

Another way of synthesising cages is by using boronic esters. Although these are mainly used for designing receptors, they are also capable of forming topologically complex molecules. Such example comes from Severin's group, which synthesised rotaxanes with boronic esters as stoppers. The reaction mixture included a crown-ether-based DN (to force the formation of [2]rotaxane), and a double bond as template for threading the DN ring.[61] The cage formed *via* this approach is pH-dependent, and can be controlled by exploiting the dynamic nature of the boronic ester. For example, no cage is obtained when using a boronic acid-based calix[3]arene along with cyclotricatechylene. However, the addition of tetraethylammonium acetate as an ion pair source has led to the self-assembly of the capsule quantitatively. This behaviour is due to the cooperative effect of the acetate; the cation acts as a template. This cage is liable to dissociation under acidic conditions, but reforms at basic pHs.[62]

Kobayashi *et al.*[63] have developed a capsule capable of binding toluene, as a receptor for detecting and separating it from reaction mixtures. The cage was formed in quantitative yield by reacting a tetraol and a triboronic acid. The addition of methanol as donor / acceptor hydrogen bonding component disrupts the cage, which re-assembles when removing the solvent from the system.

The last part of this section is a tribute to a relatively old, but important work. This comes from Jean-Marie Lehn, whose group has synthesised a circular double helicate by mixing a tribipyridine with $FeCl_2$.[64] The structural unit can be described as a torus, in which every Fe^{II} is surrounded by three different bipyridine moieties. This yields a ring composed of five tribipyridines, five metal ions, and a chloride ion in the central cavity. By replacing the metal ion (Fe^{II}) with Ni^{II}, a triple helicate $[Ni_3L_3]^{6+}$ has been obtained. Both structures can be formed in a combinatorial fashion by mixing the ions and the ligand; the triple helicate is initially formed, which further converts into the circular helicate. A subsequent study has shown that the arrangement of such molecules depends on the anion: changing the anion from Cl^- to BF_4^-, and in the presence of the same ligand, a circular helicate, $[Fe_6L_6]^{12+}$, has formed.

Recently, this helicate strategy was used by Leigh's group to synthesise a pentafoil knot as well as a Solomon link.[55] The functionalisation of the same tribipyridine ligand used by Lehn's group with flexible chains connecting the bipyridine units, and subsequent ring closing metathesis have led to a knot with eight braidings (Figure 1.23).[65]

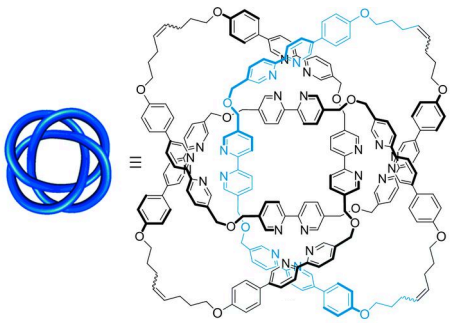


Figure 1.23. The structure of the knot with eight crossings.[65]

1.7. Mixing reversible bonds

Some systems used in DCC rely on just one type of reversible exchange and this has led to a large variety of interlocked molecules being synthesised.[66–85]

However, in most of the cases, the systems generally involve two different exchange reactions, whereas others can include up to three or four components. Despite of the arduous work invested on complex DCLs undergoing multiple exchanges, the outcome is generally trivial, leading to macrocycles. However, few groups describe the concept of associating reversible bonds to generate complex topologically molecules, and a summary of the work reported so far in the literature is provided.

The first example comes from Leigh's group, which have introduced the molecular "walker" – as a component "walking" on a track. The walker is moved along the track using disulphide and hydrazone exchanges.[66,70] To prevent the walker falling off the track, it is covalently linked either *via* S-S or C=N scaffolds. The "movements" of these molecules linked either 1-2 or 3-4 are shown in Figure 1.24. The first step consists of treating compound *1,2*-**1.17** with trifluoroacetic acid to promote hydrazone exchange, yielding a 1-2 and 2-3 linked walker. The disulphide exchange is driven by the addition of a strong base (*i.e.* DBU). To this, a disulphide initiator (DTT) is added, generating a mixture of 1-2, 2-3, 3-4 and 1-4 isomers, which is the second step. After three repetitions of this cycle, the system has reached the global minimum energy. This has been the case regardless the sequence of the cycle (*i.e.* either acid-base or base-acid), suggesting that the system tends to independently reach the minimum energy.

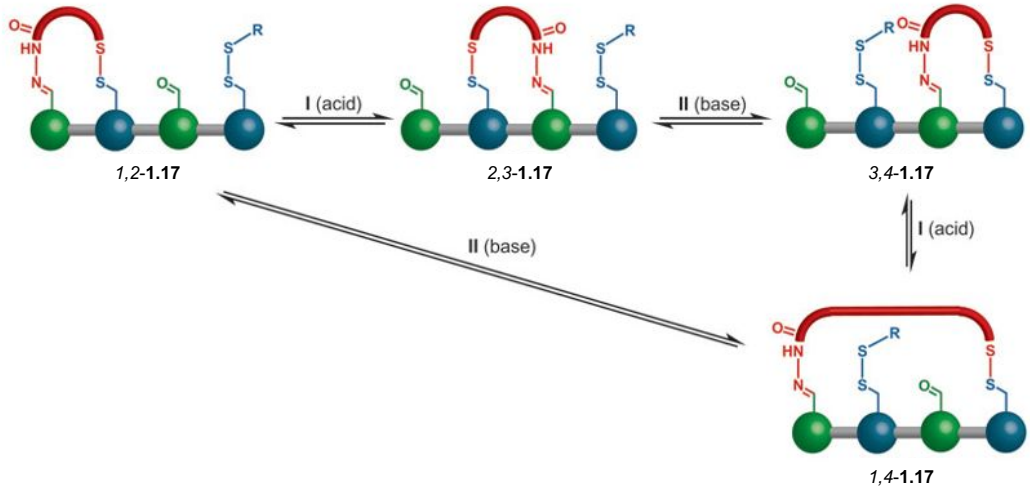


Figure 1.24. "Movements" of a molecular "walker" achieved by successive treating with acid / base.[65]

The literature also reports several capsules obtained by mixing imine and boronic ester exchange reactions. For example, Nietschke's group reported a cage formed by mixing cyclotricatechylene, *m*-xylylenediamine and 2-formylphenylboronic acid. The driving force of this reaction is the formation of a very electron-deficient group due to binding of boron to a catecholate unit. The entire system is stabilised by a dative bond between nitrogen and boron.[86] This reversible system has been also used by Severin's group, who used a [6+3+2] condensation reaction between 4-formylphenylboronic acid, pentaerythritol and tris(2-aminoethyl)amine) to form a capsule in 82% yield (Figure 1.25).[87]

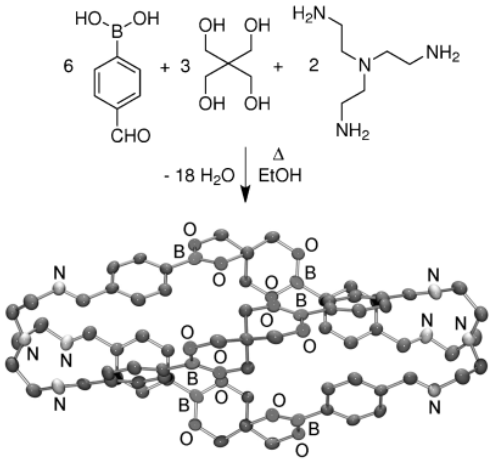


Figure 1.25. One-step synthesis of the capsule reported by Severin's group *via* [6+3+2] condensation.[87]

Another representative example obtained *via* this method is a molecular robotic arm able to transport a "cargo" along two sites, reported by Leigh's group.[29] This includes a platform with two receiving sites, and an arm covalently linked and hydrogen-bonded to the platform that enables its movement along the sites. A hydrazine-functionalised "cargo" is attached at one end, and a thiol at the other end for cargo-arm linkage. For clear distinction between the two sites, one of them is deuterium labelled. Providing that the entire molecule contains a mixture of *Z* and *E* forms in 7:3 ratio, the first step is to position the arm next to the "cargo" using TFA. Upon oxidation of thiol to disulphide with iodine, the arm, the "cargo" and the platform link together. Subsequent addition of 70 equivalents of TFA promotes the hydrazone exchange and the *Z*/*E* isomerisation process takes place, moving the "cargo" from one site to another. The reduction with tris(2-carboxyethyl)phosphine (TCEP) allowed the

arm to be detached from the “cargo”, moving it on the other side of platform. The synthetic procedure also works if started from the opposite direction (Figure 1.26).

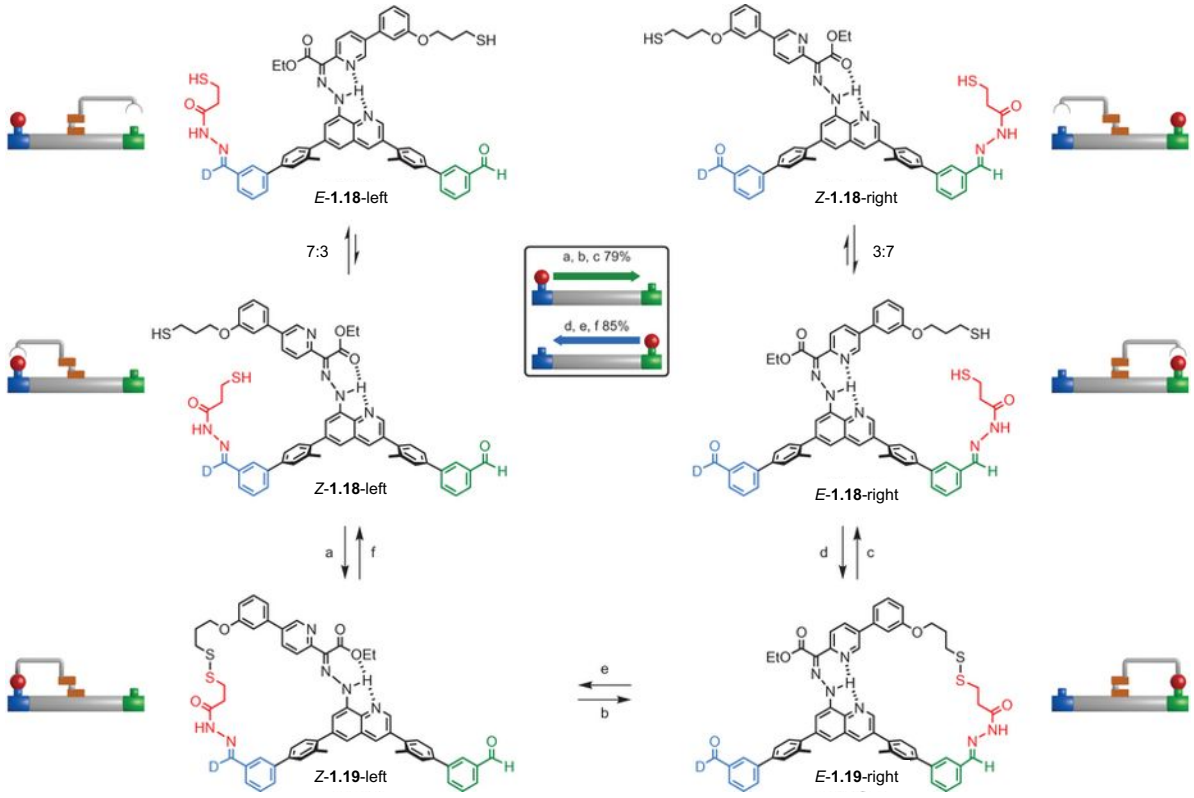


Figure 1.26. The step required for the molecular arm to transport the “cargo” from one side to another. Steps a-f are the conditions required for the transformation (these are not listed because they are not the subject of this review). 7:3 and 3:7 represent the ratio between *E* and *Z* forms.[29]

1.8. Other interlocked molecules

Another reversible bond of importance for the DCC field is hydrazone exchange; its mechanism is similar to imine exchange. In the previous example (Figure 1.24), Leigh’s group used it in conjunction to disulphide exchange and hydrogen bonding.[29]

Recently, Cougnon’s group designed a system based on hydrazone exchange, yielding a [2]catenane, a Solomon link and a trefoil knot.[88,89] The same dialdehyde building block was used for all libraries, varying only the dihydrazine component. The use of various aliphatic or aromatic hydrazines has led to different interlocked molecules. It has been shown that the structure of the Solomon link formed in these libraries is solvent-dependent. The ^{1}H NMR spectra in deuterated acetonitrile and

water suggested two different conformations. Mechanistic studies have also shown that the structure has a binding pocket for water (Figure 1.27).

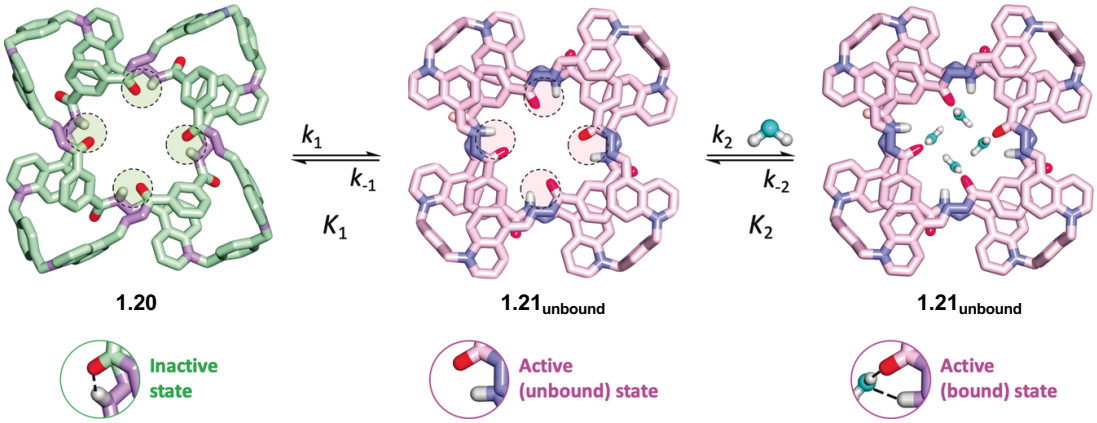


Figure 1.27. The proposed mechanism for switching between the conformers (**1.20** and **1.21**) of the Solomon link.[88]

Another strategy for making interlocked molecules is the so-called "templated-directed synthesis". In this case, the ligands assemble around a metal centre following the closing of the loop. A simple search will return thousands of entries, but just few representative examples will be described herein. The first one comes from Au-Yeung's group in Hong Kong, which have synthesised [3], [4] and [6]catenanes. Their method combines the advantages of metal-templation and the cucurbit[6]uril catalysed alkyine-azide cycloaddition.[90]

Leigh's group reported the use of the helicate approach followed by ring closure to yield knots with eight[65] and nine crossings,[91] (the later one has the highest possible number of crosses). Besides these, the pentafoil knot,[55] molecular and granny knot[92] were also synthesised. Links were obtained using similar strategies, worth mentioning being the 9_7^3 and 6_3^2 links.[91,93]

The last example comes from Stoddart, whose group developed all-acceptor and all-donor catenanes. In the first case, they assembled a radical [2]catenane from cyclobis(paraquat-*p*-phenylene) followed by its reduction to yield an octacation [2]catenane, characterised by ^{1}H NMR spectroscopy. Using either chemical and / or electrochemical approaches, different redox states were obtained, going from the tetracationic tetraradical species to the octacation one. Recently, the tetracosacataionic radial [5]catenane **1.22** has been obtained, which, as the authors stated, can be incorporated into redox active materials (Figure 1.28a).[94,95] For the

synthesis of the all-donor catenane, the same group used a similar strategy. Firstly, a donor-acceptor [2]catenane was synthesised based on 1,5-dioxynaphthalene as π-donor and the same cyclobis(paraquat-*p*-phenylene) as π-acceptor. In this case, the key step has been the reduction of the π-acceptor to its fully neutral form, which becomes a π-donor even when aromaticity is lost (Figure 1.28b).[96]

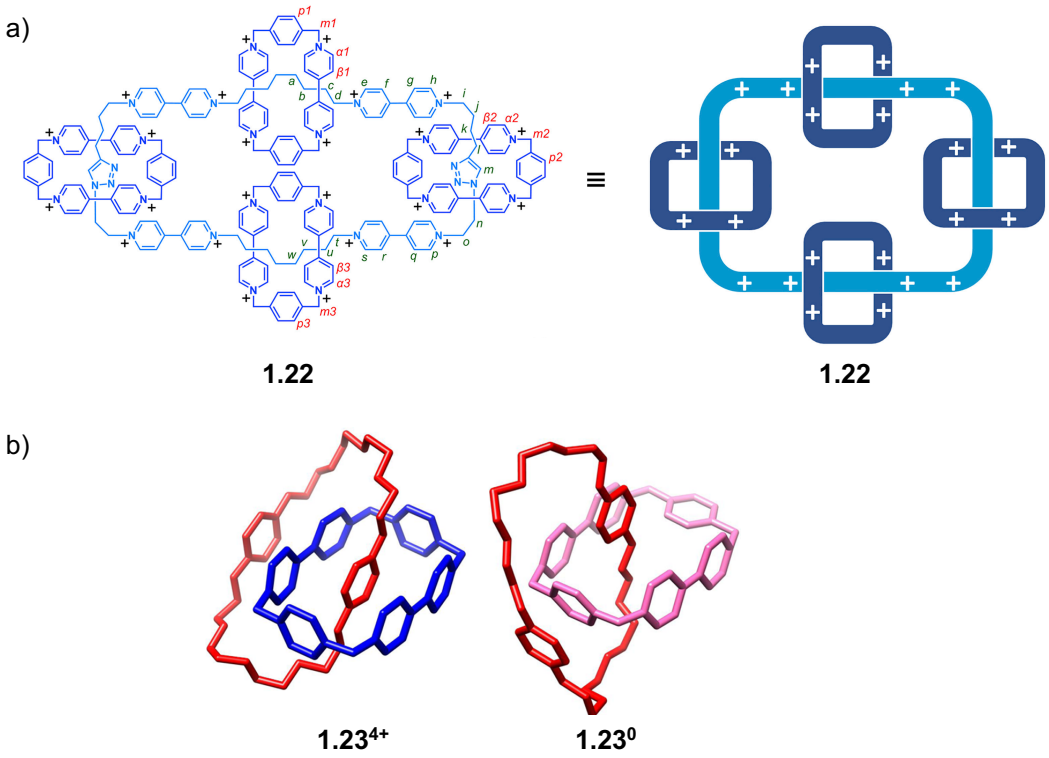


Figure 1.28. a) The structure of the tetracosacataionic radial [5]catenane **1.22**, and b) the crystal structure of the aromatic tetracataionic [2]catenane **1.23**$^{4+}$ and its reduced form **1.23**^{0} (one of the ring lost its aromaticity).[95,96]

1.9. Structurally Divergent Reaction on a Racemic Mixture (SDRRM)

All the examples described in this review (and all the others not presented herein) show the versatile and powerful nature of DCC to synthesise a large variety of interlocked molecules. DCC approach occupies a key position in the world of supramolecular chemistry, however, all the DCC-related examples reported so far in the literature are not enantioselective. This idea is explored in Chapter 2, thus the concept of divergent reaction on racemic mixture is introduced herein.

In modern organic chemistry, it is crucial to obtain compounds that are either enantio-enriched or diastereo-enriched. So far, this is mainly achieved *via* three

different routes: simple kinetic resolution, parallel kinetic resolution, and divergent reaction on a racemic mixture (Figure 1.29).[97]

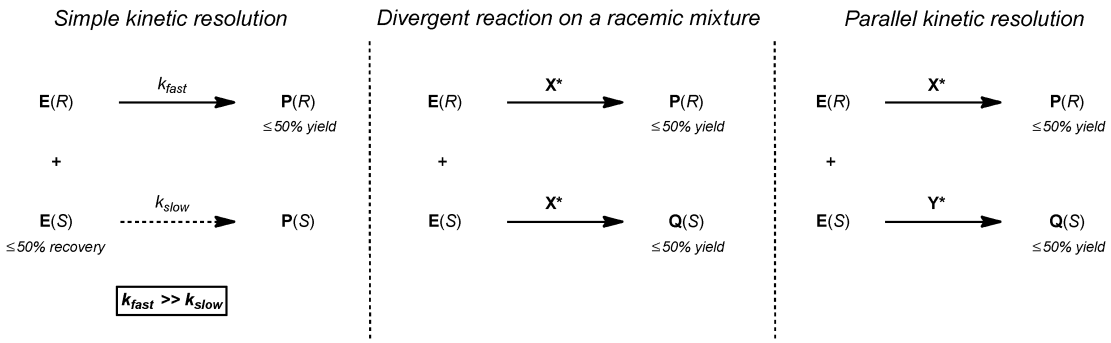


Figure 1.29. The schematic representation of simple kinetic resolution (left), divergent reaction of a racemic mixture (centre) and parallel kinetic resolution (right).[97]

The simple kinetic resolution relies on one enantiomer reacting faster with the subtract than its counterpart. If this is not the case, the reaction will produce a racemic mixture. The maximum yield in this case can be 50%, with the recovery of the unreacted enantiomer. In dynamic kinetic resolution, the slow reacting enantiomer is converted to the one reacting faster. In parallel kinetic resolution, both enantiomers are reacting at different rates with chiral molecules, giving rise to diastereoisomers, with the maximum yield 50%. Among the reaction types shown in Figure 1.29, the divergent reaction is the rarest, with only few examples reported so far. In this case, both enantiomers react with the same chiral compound, yielding a mixture of structurally divergent compounds. To the best of our knowledge, only four Structurally Divergent Reactions on a Racemic Mixture (**SDRRMs**)[98–103] have been reported in the literature (Figure 1.30); out of these, just two[99,100] lead to non-isomeric structures.

The structural divergence was usually achieved through the ability of one of the diastereomeric products to undergo a rearrangement or participate in a subsequent reaction with another compound within the reaction mixture. Without exception, these four known SDRRM examples have been developed in organic solvents and relied on structural constraints imposed by a bulky chiral catalyst.

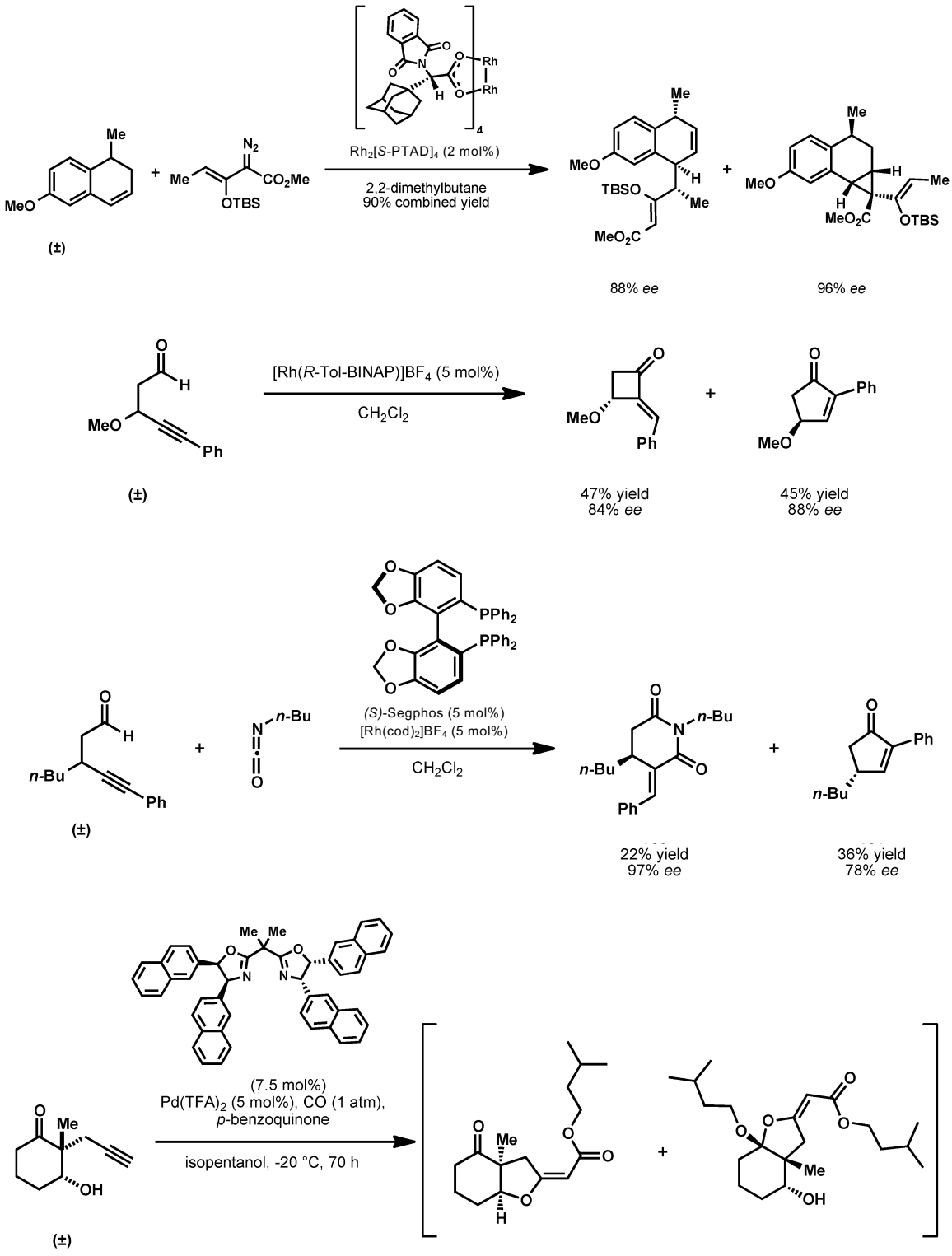


Figure 1.30. The four types of SDRRM reported so far in the literature.[98–101]

1.10. Conclusion

This short review presents few representative examples of the vast work reported in the field of DCC and supramolecular chemistry. The main methods behind the formation of these interlocked molecules and the interactions governing the assemblies are also described. There is a general increasing interest for using DCC and supramolecular chemistry-related methods due to their wide applicability from receptors design to (nano)materials. The large variety of building blocks and reversible exchanges available for DCC can lead to new classes of compounds. DCC is a powerful technique in generating new complex molecules, while the metal-templated approach is based on the assembly of ligands around a metal. Lastly, the structurally divergent reactions on racemic mixtures are introduced. This will make the subject of Chapter 2.

Aim of the research project

The aim of my Ph.D. is to design new π-donor building blocks and study their behaviour in the presence of π-acceptor molecules (*i.e.* NDI - naphthalenediimide derivatives). Another goal is taking the design of building blocks one step forward and synthesise acceptor-donor-acceptor trimers and explore their potential in making interlocked molecules in aqueous systems. All the reactions will be under thermodynamic control, using disulphide exchange.

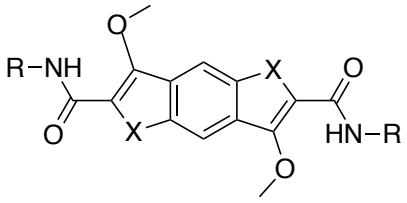


Figure 1.31. The general structure of the tri-fused heterocyclic system.

The new π-donor building blocks need to have large aromatic surfaces with a negative Q_{zz} that is comparable to the DN work that has previously been done in the Sanders' / Pantoș' groups. A literature survey indicates that tri-fused heterocyclic system (*e.g.* pyrroloindoles, benzodithiophene, benzodifuran) are good candidates. In the light of the new synthesis developed for pyrroloindoles[104] and the available

syntheses for the other two molecules we have decided to investigate their properties. Another structural feature we wanted to incorporate is the methoxy group (Figure 1.31), which would increase the electron richness and also will lock into place the amide making it a quasi-fused pentacyclic structure. Why do we want to explore this? Because we want to see how DCL will respond to a larger electron rich p-surface, contrasting it to the already studied DN. We believe that the folding properties of these *D-A* structures are largely influenced by the hydrophobic effect and using this approach is one way we can test this hypothesis. Viable synthetic methods will be developed for 2,6-dicarboxylates of benzodithiophene, benzodifuran and pyrroloindole. Further functionalisation of the carboxylate moiety with both cysteine enantiomers will allow us to test the influence of point chirality on supramolecular assemblies.

1.11. References

1 G. Gil-Ramírez, D. A. Leigh and A. J. Stephens, *Angew. Chem. Int. Ed.*, 2015, **54**, 6110–6150.

2 S. P. Black, J. K. M. Sanders and A. R. Stefankiewicz, *Chem. Soc. Rev.*, 2014, **43**, 1861–1872.

3 P. T. Corbett, J. Leclaire, L. Vial, K. R. West, J.-L. Wietor, J. K. Sanders and S. Otto, *Chem. Rev.*, 2006, **106**, 3652–3711.

4 F. B. L. Cougnon and J. K. M. Sanders, *Acc. Chem. Res.*, 2012, **45**, 2211–2221.

5 S. Otto, *Chem. Sci.*, 2013, **4**, 2953–2959.

6 D. Fass and C. Thorpe, *Chem. Rev.*, 2018, **118**, 1169–1198.

7 R. W. Visschers and H. H. J. de Jongh, *Biotechnol. Adv.*, 2005, **23**, 75–80.

8 T. Liu, Y. Wang, X. Luo, J. Li, S. A. Reed, H. Xiao, T. S. Young and P. G. Schultz, *Proc. Natl. Acad. Sci.*, 2016, **113**, 5910–5915.

9 I. Bosnjak, V. Bojovic, T. Segvic-Bubic and A. Bielen, *Protein Eng. Des. Sel.*, 2014, **27**, 65–72.

10 M. Qin, W. Wang and D. Thirumalai, *Proc. Natl. Acad. Sci.*, 2015, **112**, 11241–11246.

11 C. A. Hunter and J. K. M. Sanders, *J. Am. Chem. Soc.*, 1990, **112**, 5525–5534.

12 R. M. Parrish and C. D. Sherrill, *J. Am. Chem. Soc.*, 2014, **136**, 17386–17389.

13 M. S. Cubberley and B. L. Iverson, *J. Am. Chem. Soc.*, 2001, **123**, 7560–7563.

14 C. R. Martinez and B. L. Iverson, *Chem. Sci.*, 2012, **3**, 2191–2201.

15 K. I. Assaf and W. M. Nau, *Angew. Chem. Int. Ed.*, 2018, **57**, 13968–13981.

16 F. B. L. Cougnon, H. Y. Au-Yeung, G. D. Pantoş and J. K. M. Sanders, *J. Am. Chem. Soc.*, 2011, **133**, 3198–3207.

17 F. B. L. Cougnon, N. Ponnuswamy, N. A. Jenkins, G. D. Pantoş and J. K. M. Sanders, *J. Am. Chem. Soc.*, 2012, **134**, 19129–19135.

18 M. E. Dehkordi, V. Luxami and G. D. Pantoş, *J. Org. Chem.*, 2018, **83**, 11654–11660.

19 N. Ponnuswamy, F. B. Cougnon, G. D. Pantoş and J. K. Sanders, *J. Am. Chem. Soc.*, 2014, **136**, 8243–8251.

20 N. Ponnuswamy, F. B. L. Cougnon, J. M. Clough, G. D. Pantoş and J. K. M. Sanders, *Science*, 2012, **338**, 783–785.

21 A. R. Stefankiewicz, M. R. Sambrook and J. K. M. Sanders, *Chem. Sci.*, 2012, **3**, 2326.

22 A. R. Stefankiewicz and J. K. M. Sanders, *Chem. Commun.*, 2013, **49**, 5820.

23 J. M. A. Carnall, C. A. Waudby, A. M. Belenguer, M. C. A. Stuart, X. Yang, J. J.-P. Peyralans and S. Otto, *Science*, 2010, **327**, 1502–1506.

24 M. Malakoutikhah, J. J.-P. Peyralans, M. Colomb-Delsuc, H. Fanlo-Virgós, M. C. A. Stuart and S. Otto, *J. Am. Chem. Soc.*, 2013, **135**, 18406–18417.

25 J. Li, P. Nowak and S. Otto, *J. Am. Chem. Soc.*, 2013, **135**, 9222–9239.

26 J. W. Sadownik, E. Mattia, P. Nowak and S. Otto, *Nat. Chem.*, 2016, **8**, 264–269.

27 L. I. James, J. E. Beaver, N. W. Rice and M. L. Waters, *J. Am. Chem. Soc.*, 2013, **135**, 6450–6455.

28 Y. Furusho, T. Oku, T. Hasegawa, A. Tsuboi, N. Kihara and T. Takata, *Chem. - Eur. J.*, 2003, **9**, 2895–2903.

29 S. Kassem, A. T. L. Lee, D. A. Leigh, A. Markevicius and J. Solà, *Nat. Chem.*, 2015, **8**, 138–143.

30 C. Tsiamantas, X. de Hatten, C. Douat, B. Kauffmann, V. Maurizot, H. Ihara, M. Takafuji, N. Metzler-Nolte and I. Huc, *Angew. Chem. Int. Ed.*, 2016, **55**, 6848–6852.

31 J. Sun, B. O. Patrick and J. C. Sherman, *Tetrahedron*, 2009, **65**, 7296–7302.

32 J. W. Sadownik, E. Mattia, P. Nowak and S. Otto, *Nat. Chem.*, 2016, **8**, 264–269.

33 A. Bugaut, K. Jantos, J.-L. Wietor, R. Rodriguez, J. K. M. Sanders and S. Balasubramanian, *Angew. Chem. Int. Ed.*, 2008, **47**, 2677–2680.

34 K. S. Chichak, S. J. Cantrill, A. R. Pease, S.-H. Chiu, G. W. V. Cave, J. L. Atwood and J. F. Stoddart, *Science*, **304**, 1318–1312.

35 C. D. Meyer, R. S. Forgan, K. S. Chichak, A. J. Peters, N. Tangchaivang, G. W. V. Cave, S. I. Khan, S. J. Cantrill and J. F. Stoddart, *Chem. - Eur. J.*, 2010, **16**, 12570–12581.

36 C. D. Pentecost, K. S. Chichak, A. J. Peters, G. W. V. Cave, S. J. Cantrill and J. F. Stoddart, *Angew. Chem. Int. Ed.*, 2007, **46**, 218–222.

37 A.-J. Avestro, D. M. Gardner, N. A. Vermeulen, E. A. Wilson, S. T. Schneebeli, A. C. Whalley, M. E. Belowich, R. Carmieli, M. R. Wasielewski and J. F. Stoddart, *Angew. Chem. Int. Ed.*, 2014, **53**, 4442–4449.

38 M. E. Belowich, C. Valente and J. F. Stoddart, *Angew. Chem. Int. Ed.*, 2010, **49**, 7208–7212.

39 R. A. Bilbeisi, T. K. Ronson and J. R. Nitschke, *Angew. Chem. Int. Ed.*, 2013, **52**, 9027–9030.

40 S. P. Black, A. R. Stefankiewicz, M. M. J. Smulders, D. Sattler, C. A. Schalley, J. R. Nitschke and J. K. M. Sanders, *Angew. Chem. Int. Ed.*, 2013, **52**, 5749–5752.

41 S. P. Black, D. M. Wood, F. B. Schwarz, T. K. Ronson, J. J. Holstein, A. R. Stefankiewicz, C. A. Schalley, Jeremy. K. M. Sanders and J. R. Nitschke, *Chem Sci*, 2016, **7**, 2614–2620.

42 M. Kieffer, B. S. Pilgrim, T. K. Ronson, D. A. Roberts, M. Aleksanyan and J. R. Nitschke, *J. Am. Chem. Soc.*, 2016, **138**, 6813–6821.

43 S. M. Jansze, G. Cecot, M. D. Wise, K. O. Zhurov, T. K. Ronson, A. M. Castilla, A. Finelli, P. Pattison, E. Solari, R. Scopelliti, G. E. Zelinskii, A. V. Vologzhanina, Y. Z. Voloshin, J. R. Nitschke and K. Severin, *J. Am. Chem. Soc.*, 2016, **138**, 2046–2054.

44 T. Hasell, X. Wu, J. T. A. Jones, J. Bacsa, A. Steiner, T. Mitra, A. Trewin, D. J. Adams and A. I. Cooper, *Nat. Chem.*, 2010, **2**, 750–755.

45 T. Tozawa, J. T. A. Jones, S. I. Swamy, S. Jiang, D. J. Adams, S. Shakespeare, R. Clowes, D. Bradshaw, T. Hasell, S. Y. Chong, C. Tang, S.

Thompson, J. Parker, A. Trewin, J. Bacsa, A. M. Z. Slawin, A. Steiner and A. I. Cooper, *Nat. Mater.*, 2009, **8**, 973–978.

46 E. A. Wilson, N. A. Vermeulen, P. R. McGonigal, A.-J. Avestro, A. A. Sarjeant, C. L. Stern and J. F. Stoddart, *Chem. Commun.*, 2014, **50**, 9665.

47 O. A. Bozdemir, G. Barin, M. E. Belowich, A. N. Basuray, F. Beuerle and J. F. Stoddart, *Chem. Commun.*, 2012, **48**, 10401–10403.

48 A. G. Slater, M. A. Little, A. Pulido, S. Y. Chong, D. Holden, L. Chen, C. Morgan, X. Wu, G. Cheng, R. Clowes, M. E. Briggs, T. Hasell, K. E. Jelfs, G. M. Day and A. I. Cooper, *Nat. Chem.*, 2017, **9**, 17-25.

49 T. Hasell, M. Schmidtmann and A. I. Cooper, *J. Am. Chem. Soc.*, 2011, **133**, 14920–14923.

50 L. Chen, P. S. Reiss, S. Y. Chong, D. Holden, K. E. Jelfs, T. Hasell, M. A. Little, A. Kewley, M. E. Briggs, A. Stephenson, K. M. Thomas, J. A. Armstrong, J. Bell, J. Busto, R. Noel, J. Liu, D. M. Strachan, P. K. Thallapally and A. I. Cooper, *Nat. Mater.*, 2014, **13**, 954–960.

51 X. Liu and R. Warmuth, *J. Am. Chem. Soc.*, 2006, **128**, 14120–14127.

52 P. Skowronek, B. Warżajtis, U. Rychlewska and J. Gawroński, *Chem. Commun.*, 2013, **49**, 2524.

53 D. M. Wood, W. Meng, T. K. Ronson, A. R. Stefankiewicz, J. K. M. Sanders and J. R. Nitschke, *Angew. Chem. Int. Ed.*, 2015, **54**, 3988–3992.

54 C. S. Wood, T. K. Ronson, A. M. Belenguer, J. J. Holstein and J. R. Nitschke, *Nat. Chem.*, 2015, **7**, 354–358.

55 J.-F. Ayme, J. E. Beves, D. A. Leigh, R. T. McBurney, K. Rissanen and D. Schultz, *Nat. Chem.*, 2012, **4**, 15–20.

56 T. Prakasam, M. Lusi, M. Elhabiri, C. Platas-Iglesias, J.-C. Olsen, Z. Asfari, S. Cianférani-Sanglier, F. Debaene, L. J. Charbonnière and A. Trabolsi, *Angew. Chem. Int. Ed.*, 2013, **52**, 9956–9960.

57 G. D. Pantoş, P. Pengo and J. K. M. Sanders, *Angew. Chem. Int. Ed.*, 2007, **46**, 194–197.

58 G. D. Pantoş, J. L. Wietor and J. K. M. Sanders, *Angew. Chem. Int. Ed.*, 2007, **46**, 2238–2240.

59 J.-L. Wietor, G. D. Pantoş and J. K. M. Sanders, *Angew. Chem. Int. Ed.*, 2008, **47**, 2689–2692.

60 C. Zhang, Q. Wang, H. Long and W. Zhang, *J. Am. Chem. Soc.*, 2011, **133**, 20995–21001.

61 N. Christinat, R. Scopelliti and K. Severin, *Chem. Commun.*, 2008, 3660.

62 K. Kataoka, T. D. James and Y. Kubo, *J. Am. Chem. Soc.*, 2007, **129**, 15126–15127.

63 N. Nishimura and K. Kobayashi, *Angew. Chem. Int. Ed.*, 2008, **47**, 6255–6258.

64 B. Hasenknopf, J.-M. Lehn, B. O. Kneisel, G. Baum and D. Fenske, *Angew. Chem. Int. Ed.*, 1996, **35**, 1838–1840.

65 J. J. Danon, A. Krüger, D. A. Leigh, J.-F. Lemonnier, A. J. Stephens, I. J. Vitorica-Yrezabal and S. L. Woltering, *Science*, 2017, **355**, 159–162.

66 M. von Delius, E. M. Geertsema and D. A. Leigh, *Nat. Chem.*, 2010, **2**, 96–101.

67 K.-D. Zhang and S. Matile, *Angew. Chem. Int. Ed.*, 2015, **54**, 8980–8983.

68 F. Schaufelberger, L. Hu and O. Ramström, *Chem. - Eur. J.*, 2015, **21**, 9776–9783.

69 A. G. Orrillo, A. M. Escalante and R. L. E. Furlan, *Chem. Commun.*, 2008, 5298.

70 M. von Delius, E. M. Geertsema, D. A. Leigh and D.-T. D. Tang, *J. Am. Chem. Soc.*, 2010, **132**, 16134–16145.

71 R. J. Sarma, S. Otto and J. R. Nitschke, *Chem. - Eur. J.*, 2007, **13**, 9542–9546.

72 A. G. Orrillo, A. M. Escalante and R. L. E. Furlan, *Chem. - Eur. J.*, 2016, **22**, 6746–6749.

73 V. Goral, M. I. Nelen, A. V. Eliseev and J.-M. Lehn, *Proc. Natl. Acad. Sci.*, 2001, **98**, 1347–1352.

74 S. Mukherjee, William. L. A. Brooks, Y. Dai and B. S. Sumerlin, *Polym Chem*, 2016, **7**, 1971–1978.

75 P. Vongvilai and O. Ramström, *J. Am. Chem. Soc.*, 2009, **131**, 14419–14425.

76 L. You, J. S. Berman and E. V. Anslyn, *Nat. Chem.*, 2011, **3**, 943–948.

77 J. Leclaire, L. Vial, S. Otto and J. K. M. Sanders, *Chem. Commun.*, 2005, 1959–1961.

78 R. C. Lirag and O. Š. Miljanić, *Chem. Commun.*, 2014, **50**, 9401–9404.

79 H. M. Seifert, K. Ramirez Trejo and E. V. Anslyn, *J. Am. Chem. Soc.*, 2016, **138**, 10916–10924.

80 K. D. Okochi, Y. Jin and W. Zhang, *Chem. Commun.*, 2013, **49**, 4418–4420.

81 M. E. Bracchi and D. A. Fulton, *Chem. Commun.*, 2015, **51**, 11052–11055.

82 Z. Rodriguez-Docampo and S. Otto, *Chem. Commun.*, 2008, 5301–5303.

83 A. M. Escalante, A. G. Orrillo and R. L. E. Furlan, *J. Comb. Chem.*, 2010, **12**, 410–413.

84 L. Rocard, A. Berezin, F. De Leo and D. Bonifazi, *Angew. Chem. Int. Ed.*, 2015, **54**, 15739–15743.

85 S. Lascano, K.-D. Zhang, R. Wehlauch, K. Gademann, N. Sakai and S. Matile, *Chem. Sci.*, 2016, **7**, 4720–4724.

86 M. Hutin, G. Bernardinelli and J. R. Nitschke, *Chem. - Eur. J.*, 2008, **14**, 4585–4593.

87 N. Christinat, R. Scopelliti and K. Severin, *Angew. Chem. Int. Ed.*, 2008, **47**, 1848–1852.

88 F. B. L. Cougnon, K. Caprice, M. Pupier, A. Bauzá and A. Frontera, *J. Am. Chem. Soc.*, 2018, **140**, 12442–12450.

89 K. Caprice, M. Pupier, A. Bauzá, A. Frontera and F. B. L. Cougnon, *Angew. Chem. Int. Ed.*, 2019, **58**, 8053–8057.

90 K. Wang, C.-C. Yee and H. Y. Au-Yeung, *Chem. Sci.*, 2016, **7**, 2787–2792.

91 L. Zhang, A. J. Stephens, A. L. Nussbaumer, J.-F. Lemonnier, P. Jurček, I. J. Vitorica-Yrezabal and D. A. Leigh, *Nat. Chem.*, 2018, **10**, 1083–1088.

92 D. A. Leigh, L. Pirvu and F. Schaufelberger, *J. Am. Chem. Soc.*, 2019, **141**, 6054–6059.

93 J. J. Danon, D. A. Leigh, S. Pisano, A. Valero and I. J. Vitorica-Yrezabal, *Angew. Chem. Int. Ed.*, 2018, **57**, 13833–13837.

94 J. C. Barnes, A. C. Fahrenbach, D. Cao, S. M. Dyar, M. Frasconi, M. A. Giesener, D. Benitez, E. Tkatchouk, O. Chernyashevskyy, W. H. Shin, H. Li, S. Sampath, C. L. Stern, A. A. Sarjeant, K. J. Hartlieb, Z. Liu, R. Carmieli, Y. Y. Botros, J. W. Choi, A. M. Z. Slawin, J. B. Ketterson, M. R. Wasielewski, W. A. Goddard and J. F. Stoddart, *Science*, 2013, **339**, 429–433.

95 M. T. Nguyen, D. P. Ferris, C. Pezzato, Y. Wang and J. F. Stoddart, *Chem*, 2018, **4**, 2329–2344.

96 F. Stoddart, W. A. Goddardlll and I. R. Fernando, *J. Am. Chem. Soc.*, 2016, 10214–10225.

97 L. C. Miller and R. Sarpong, *Chem. Soc. Rev.*, 2011, **40**, 4550–4562.

98 E. Nadeau, D. L. Ventura, J. A. Brekan and H. M. L. Davies, *J. Org. Chem.*, 2010, **75**, 1927–1939.

99 K. Tanaka, Y. Hagiwara and M. Hirano, *Angew. Chem. Int. Ed.*, 2006, **45**, 2734–2737.

100 K. Kato, S. Motodate, S. Takaishi, T. Kusakabe and H. Akita, *Tetrahedron*, 2008, **64**, 4627–4636.

101 K. Tanaka and G. C. Fu, *J. Am. Chem. Soc.*, 2003, **125**, 8078–8079.

102 H. M. L. Davies and A. M. Walji, *Angew. Chem. Int. Ed.*, 2005, **44**, 1733–1735.

103 H. M. L. Davies, X. Dai and M. S. Long, *J. Am. Chem. Soc.*, 2006, **128**, 2485–2490.

104 W. L. H. Iv, C. S. Gelbaum, L. Gelbaum, P. Pollet, K. W. Richman, W. DuBay, J. D. Butler, G. Wells and C. L. Liotta, *RSC Adv.*, 2013, **3**, 13232–13242.

Chapter 2: Structural Divergent Reaction on Racemic Mixture

One of the first ideas as part of my Ph.D. was the design and synthesis of new aromatic π-donor building blocks. To further expand the world of DCC, new building blocks are needed. This would be a replacement for the much-used dialkoxynaphthalene (DN) molecules. The molecules would be more (or less) electron-rich and would have a larger π surface area than the DN. We have decided to design and synthesise a building block using a dithiophene derivative – a benzo[1,2-*b*:4,5-*b'*]dithiophene molecule (BDT), bearing cysteine appendages for disulphide exchange. Computational calculations (Geometry optimisation was performed using Avogadro[1] (Force field: UFF, Algorithm: Conjugate gradients). This was followed by MP2 optimisation), which confirmed the electron-rich property of this molecule, with a Q_{zz} of -23 B. Thus, the new building block would be a good partner for the electron-deficient NDI (Q_{zz} of 19.5 B).

The synthetic pathway included the esterification of the dibromo derivative **2.3** (The numbering of the compounds started with **2.3**, just to avoid any confusions with *R,R*-**1** and *R,R*-**2** molecules), followed by the palladium-mediated S-Br exchange. Then the benzodithiophene cycle was closed using potassium tert-butoxide in a Claisen type mechanism (compound **2.5** from Scheme 2.1). The methylation of the free phenolic groups and the hydrolysis of the ester group yielded the carboxylic acid derivative **2.7**. NHS activation of the carboxylic group **2.8** and further reaction with *S*-trityl-L-cysteine afforded the BDT with cysteine appendages **2.9**. The last step consisted of trityl deprotection under acidic conditions to give the product that can be used in DCC (*R,R*-**2**). Single crystals of **2.9** were grown from a solution of acetone and trifluoroethanol. The X-ray crystal structure (Figure 2.1) shows, besides trityl group interconnections, an exciting feature: a hydrogen bond between the amide NH and methoxy O. This affords the system to adopt a pseudopentacyclic (three aromatic cores from benzodithiophene and two pseudocyclic structures held together by hydrogen bonding between amide and one of the methoxy groups) structure instead of three aromatic units, thus a larger π surface. Studies on similar frameworks have shown that, even in water, this hydrogen bonding is stable.[2–6]

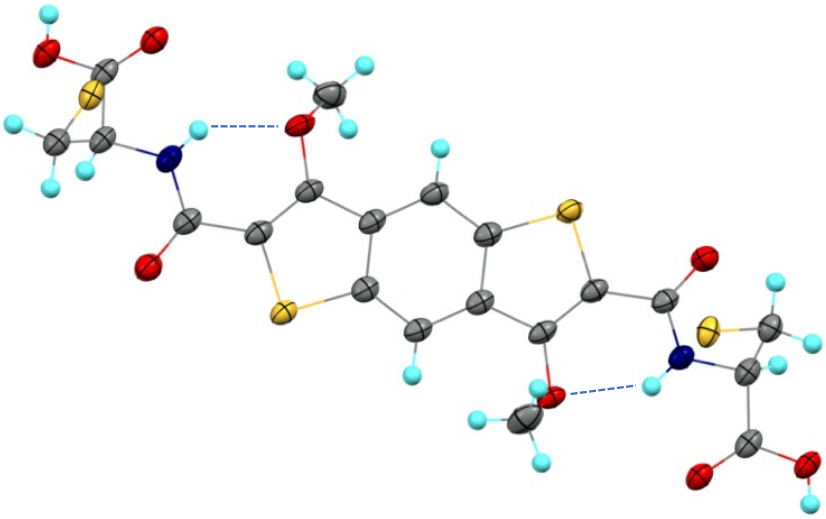
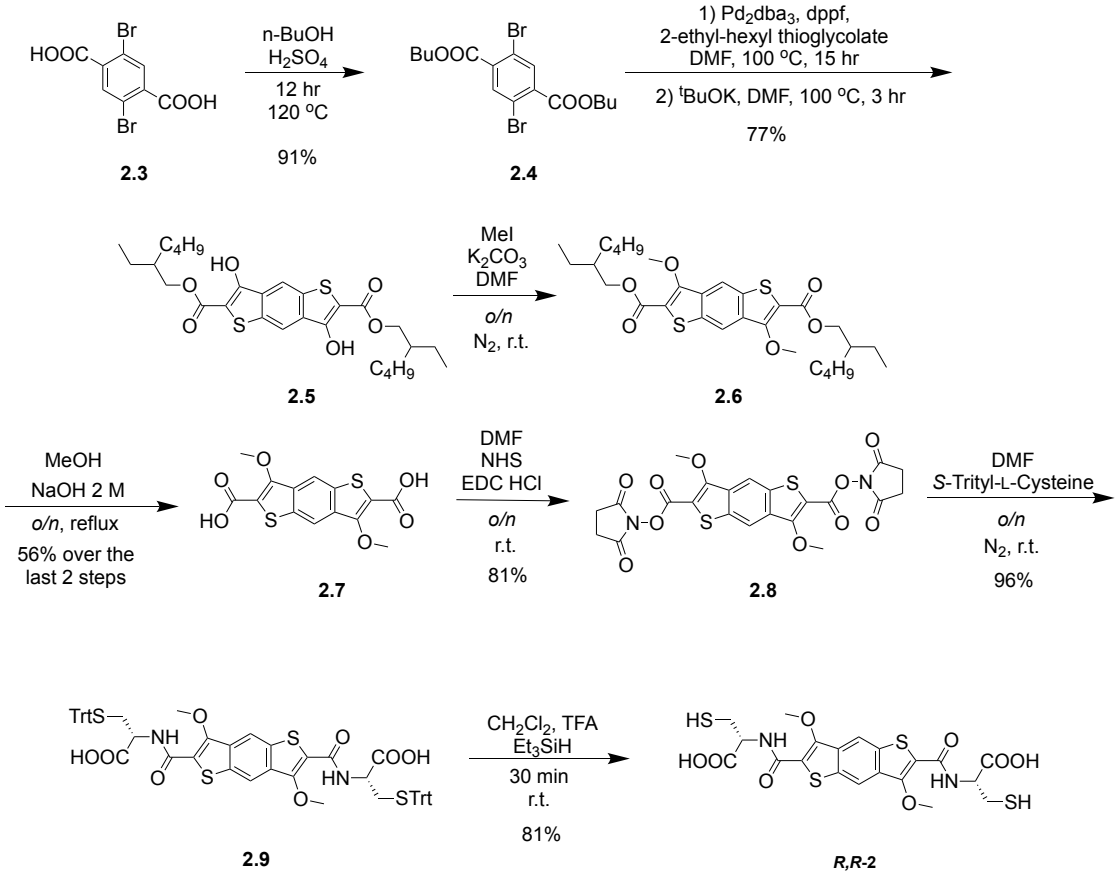
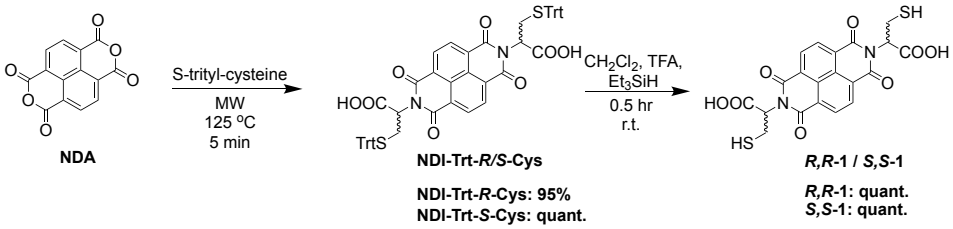


Figure 2.1. Single crystal X-ray structure of compound **2.9** precursor showing the presence of the quasi-pentacyclic core (through the hydrogen bonding (blue dashed line) between NH and O from methoxy). H atoms and the trityl groups are missing for clarity. (C – grey, H – cyan, O – red, N – blue, S – yellow).



Scheme 2.1. The synthetic scheme for *R,R*-**2**.

The enantiopure acceptor building blocks (*R,R*-**1** and *S,S*-**1**) are naphthalene-based molecules synthesised using a procedure developed in Sanders' and Pantoș' groups. They also have cysteine attached to allow for disulphide exchange (Scheme 2.2).[7,8]



Scheme 2.2. The synthetic scheme for *R,R*-**1** and *S,S*-**1**.

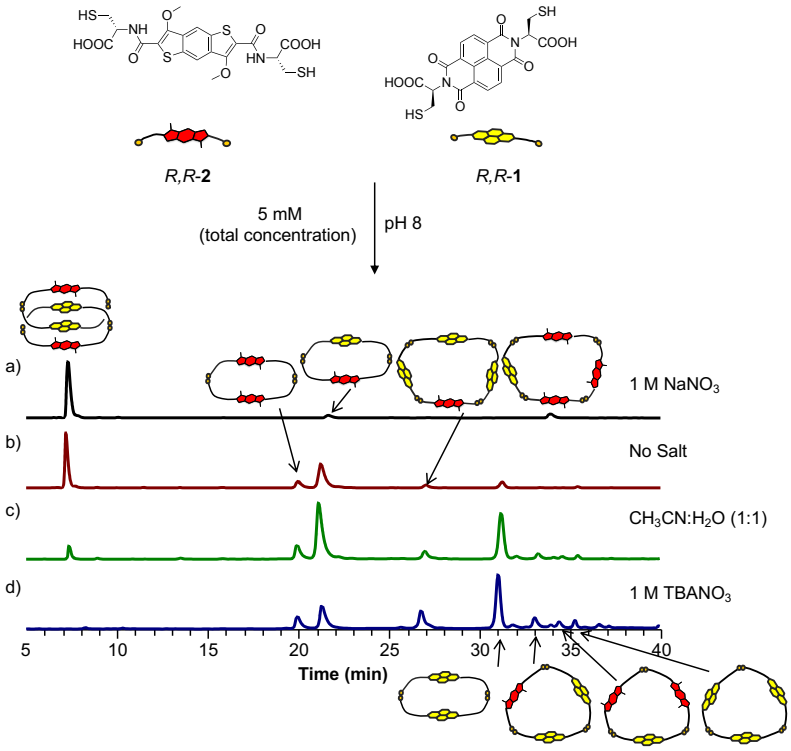
Firstly, the library outcome between *R,R*-**2** and the *R* isomer of NDI (*R,R*-**1**) will be discussed followed by the heterochiral library result (between *R,R*-**2** and the *S* isomer of NDI (*S,S*-**1**)). The discussion will then focus on the behaviour of the NDI racemic mixture (throughout the thesis a racemic mixture is a 1:1 mixture of enantiomers produced from gravimetric measurement of the quantities required of each enantiomer) and *R,R*-**2** DCLs.

2.1. Analysis of *R,R*-1 and *R,R*-2 library

The dynamic combinatorial libraries (DCLs) were set by dissolving the building blocks in 10 mM NaOH to a total concentration of 5 mM. The pH was then adjusted to 8 using 100 mM NaOH / 100 mM HCl. The libraries (in the presence or absence of additives, or with / without competitors) were stirred in air-capped vials for at least three days to allow for full oxidation of the thiols.

When the building blocks were mixed in 1:1 ratio without any additives, a mixture of dimers (hetero and homodimer, 64%), tetramers (8%) and a [2]catenane (**Cat I *RRRR***) (32%) were formed. The short retention time and the MS and MS/MS analyses indicated the [2]catenane contains two *R,R*-**1** and two *R,R*-**2** units (Figures 2.4 and 2.5). In an attempt to increase the yield of the catenane, the library was repeated in the presence of 1 M $NaNO_3$, which is known for enhancing the hydrophobic effect.[8] The yield of the catenane has indeed increased up to 76%, which supports that the hydrophobic effect drives the formation of this [2]catenane. When

the hydrophobic effect was decreased by setting up the library in a water : acetonitrile mixture (1:1) (because acetonitrile is less polar than water), the proportion of catenane was only 3%. The last library contained 1 M tetrabutylammonium nitrate ($TBANO_3$), which can solvate the aromatic cores. In this case, only macrocycles were obtained. The results are summarised in Figure 2.2 (the full characterisation of each compound can be found at the end of the chapter in Figures 2.51 – 2.58).



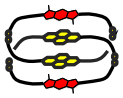
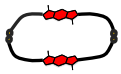
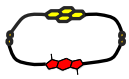
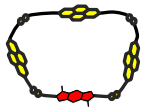
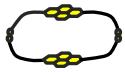
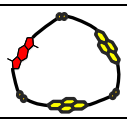
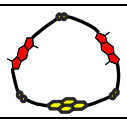
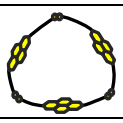
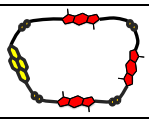
Species					
i.d.	**Cat I *RRRR***	**Y**	**X**	***AAAD* (*RRRR*) macrocycle**	***R,R*-1 dimer**
Species					
i.d.	***AAD* (*RRR*) macrocycle**	***ADD* (*RRR*) macrocycle**	***R,R*-1 trimer**	***DDDA* (*RRRR*) macrocycle**	

Figure 2.2. Reverse-phase HPLC analysis of ***R,R*-1**:***R,R*-2** (1:1 molar ratio, 5 mM total concentration) library a) in the presence of 1 M $NaNO_3$, b) no salt, c) in a mixture of $CH_3CN:H_2O$ (1:1) and d) in the presence of 1 M $TBANO_3$. Absorbances recorded at 389 nm. The unlabelled peaks did not ionise and could not be identified.

2.1.1. Analysis of the BDT-NDI DAAD [2]catenane

One way to identify if a molecule has an interlocked structure, besides a short retention time (oligomers which increase in size elute normally in a progression style based on their molecular weight, whereas catenanes have a shorter retention time due to stacking) on the chromatogram and 1D and 2D NMR analyses, is MS/MS fragmentation. Let us consider a tetramer and a [2]catenane, which are structural isomers, *i.e.* both have the same molecular formula. A catenated structure will have fragments with a *m/z* equal to the half of molecular weight or lower in the case of *DAAD* sequence (for singly charged molecules). On the other side, a *DADA* [2]catenane would fragment into a *DD* or *AA* dimers. This is not true for the tetramer though. In this case, there will be open trimer or species with *m/z* between dimer and tetramer .

The concept is presented in Figure 2.3:

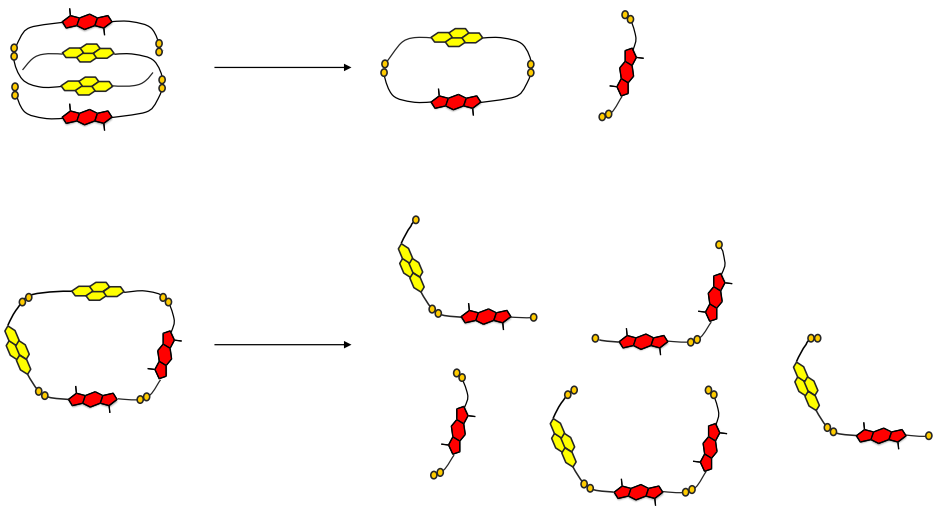


Figure 2.3. The theoretical fragmentation of a [2]catenane (top) *vs.* a tetramer (bottom).

The [2]catenane was isolated from a 3 mL library *via* semi-preparative HPLC. The chromatogram of the isolated catenane is shown in Figure 2.4.

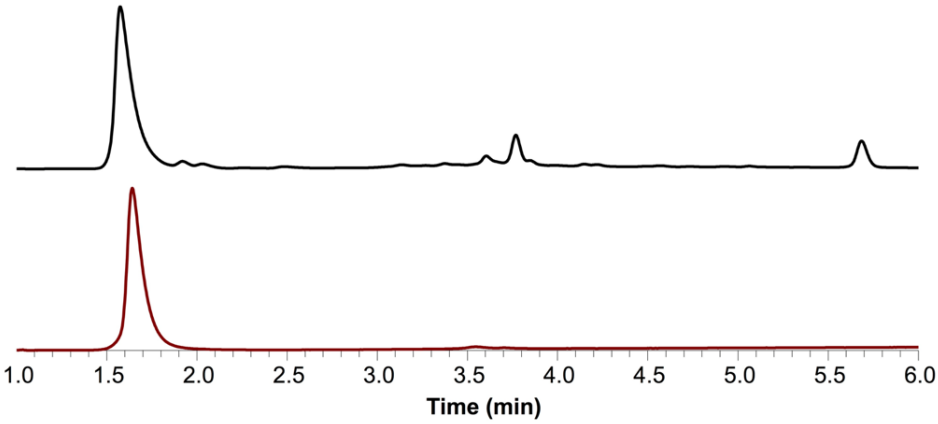


Figure 2.4. Reverse-phase HPLC analysis of *R,R*-**1**:*R,R*-**2** (1:1 molar ratio, 5 mM total concentration) library before (top trace) and **Cat I *RRRR*** after isolation (bottom trace). Absorbances recorded at 389 nm.

The isolated catenane was subjected to ^{1}H NMR and CD / UV experiments to further prove its interlocked nature.

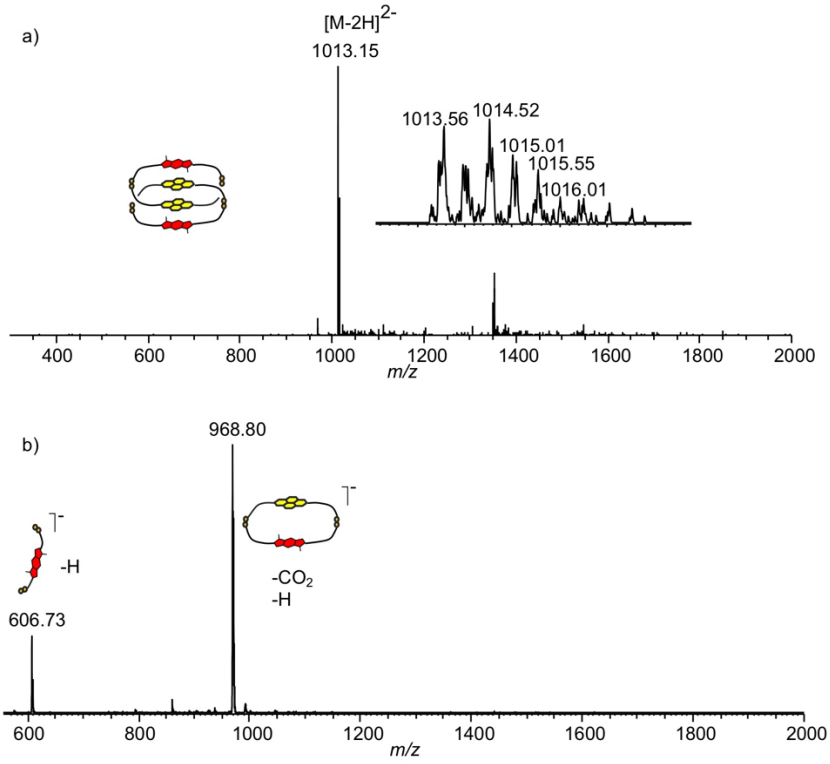


Figure 2.5. a) MS (*-ve*) of **Cat I *RRRR***; zoom of the molecular ion is shown as an inset and b) MS/MS (*-ve*) of **Cat I *RRRR***. The presence of perthiolate was reported before (the peak at *m/z* 1350 corresponds to an aggregate).[8]

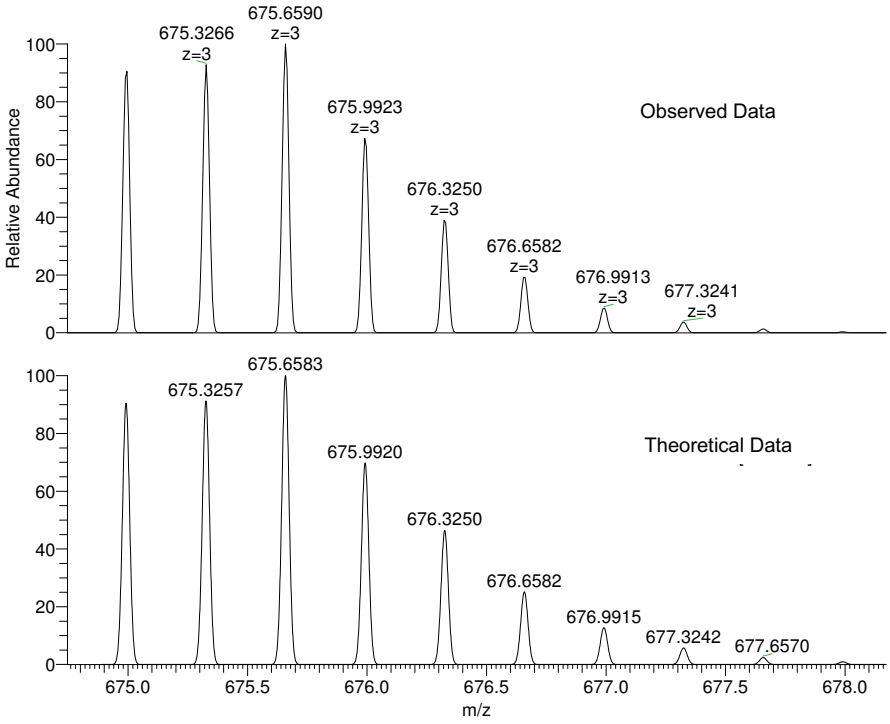


Figure 2.6. High resolution MS of **Cat I *RRRR***.

Figure 2.6 shows the MS spectrum of **Cat I *RRRR*** along with its MS/MS spectrum. There are no species with *m/z* larger than that of a heterodimer, which strengthens the assumption of the [2]catenane with a *DAAD* stacking sequence, as shown in Figure 2.5 (*D* from donor: BDT or *R,R*-**2**, and *A* from acceptor: NDI or *R,R*-**1**; the peak at *m/z* 1350 corresponds to an aggregate). It is impossible to have a *DADA* arrangement because previous studies in Sanders' group have showed that the cavity of *AA* homodimer is too small to accommodate other molecules inside, and the MS/MS would show fragments from homodimers.[9] The question that remains is thus: is it an *ADDA* or *DAAD* sequence? The NMR analysis gives the answer.

The first feature to note is the upfield shift of all of the aromatic protons due to the ring current effect of the NDI and BDT stacked cores; this is in agreement with a catenated structure. The peak at 9.23 ppm corresponds to the NH group, and its sharpness indicates it does not get exchanged with D_2O. The VT-NMR (Figure 2.7) experiment confirms that the structure is a catenane: in the case of an isomeric macrocycle, due to large thermal energy at high temperature, the ^{1}H NMR peaks will broaden. In an opposite scenario, a macrocycle would show broad peaks at room

temperature that would become sharp as the temperature increases due to fast rotation and the signals average out.

The sharpness of the peaks and the relatively constant chemical shifts as the temperature increases are supportive of an interlocked geometry for this molecule. The similarity in the chemical shift with the previous isolated [2]catenanes in Sanders' / Pantoș' groups indicates a *DAAD* stacking. The HSQC (Figure 2.9) also gives useful information about the carbon atoms as the sample is too diluted for a normal ^{13}C NMR experiment.

We also looked at the chiral properties of **Cat I *RRRR***. Figure 2.10 shows the presence of a charge transfer (CT) band (λ_{max} = 560 nm), which indicates once again that the BDT and NDI cores are in close proximity, which is due to the tightness of the system. The formation of the CT complex is indicated by a bathochromically shifted band compared to individual chromophores of each component. The circular dichroism (CD) spectrum of this catenane shows a large bisignate Cotton effect, with a positive maximum around 300 nm. The CD spectrum maintain its profile at high temperatures; minimum degradation is observed. The minimum change in CD / absorbance over 5 – 85 °C temperature range reinforces the idea of an interlocked structure (Figure 2.10).

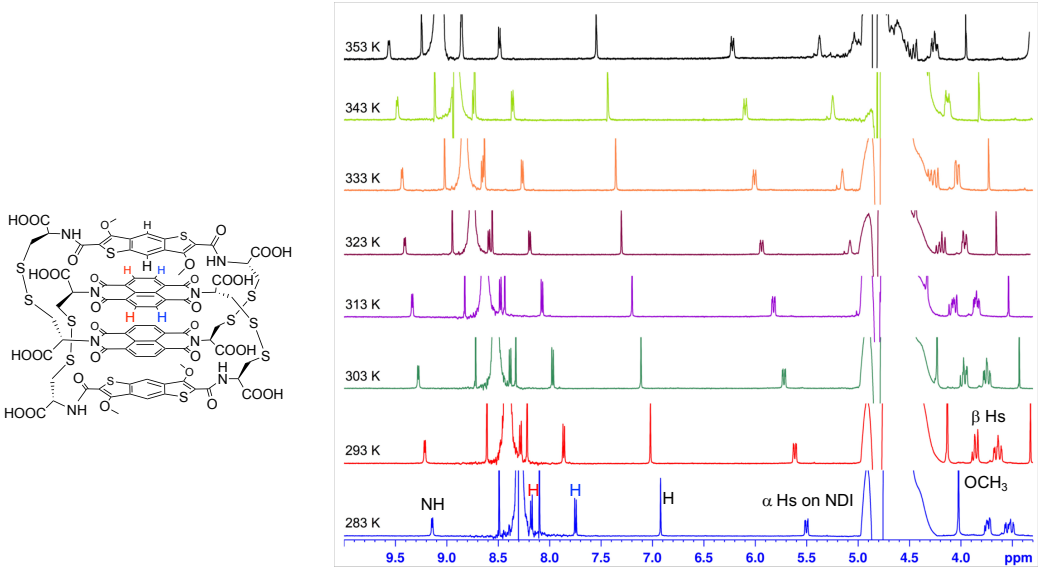


Figure 2.7. Variable temperature ^{1}H NMR spectra (500 MHz) of **Cat I *RRRR*** at the temperatures indicated above each spectrum. The residual solvent (H_2O) was referenced at 4.79 ppm. The spectrum was acquired in an H_2O:D_2O (~4:1) mixture.

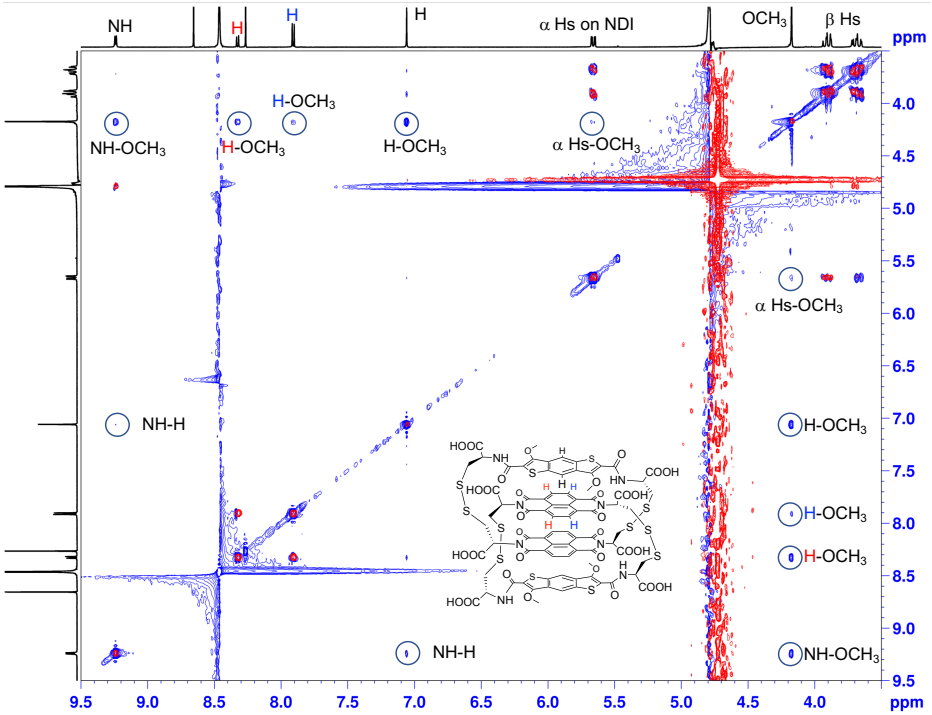


Figure 2.8. ^{1}H-^{1}H COSY (red) and 2D NOESY (blue) spectra (500 MHz, 298 K) superimposed of **Cat I *RRRR***. The residual solvent (H_2O) was referenced at 4.79 ppm. The spectrum was acquired in an H_2O:D_2O (~ 4:1) mixture.

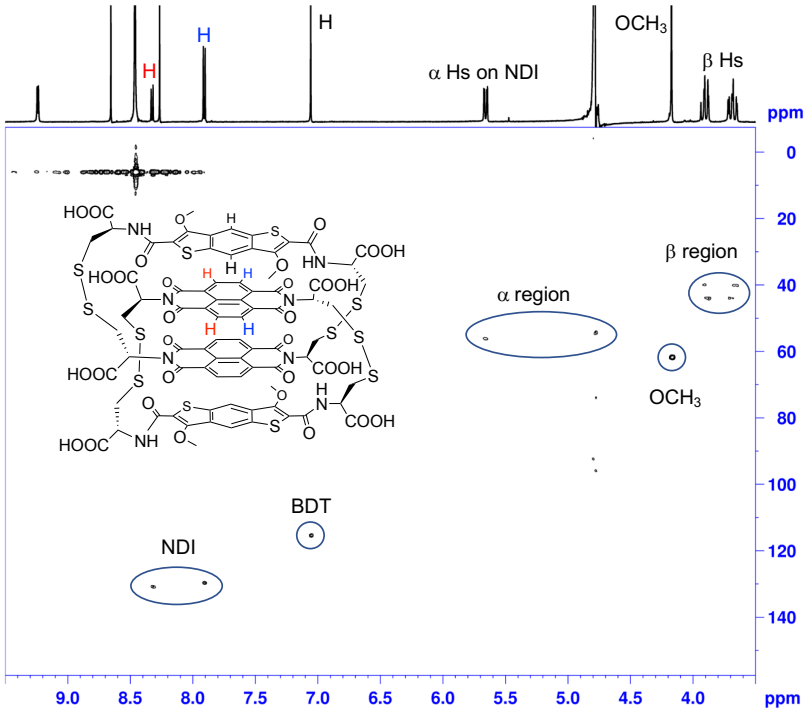


Figure 2.9. HSQC spectrum (500 MHz, 298 K) of **Cat I *RRRR***. The solvent (H_2O) was referenced at 4.79 ppm. The spectrum was acquired in an H_2O:D_2O (~ 4:1) mixture.

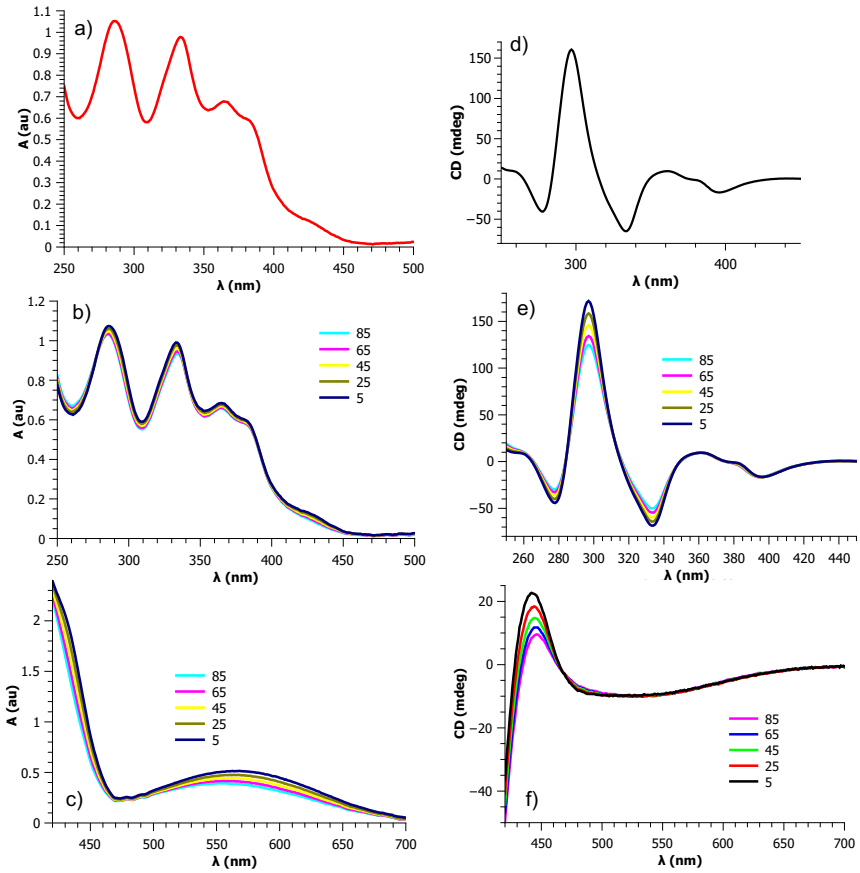
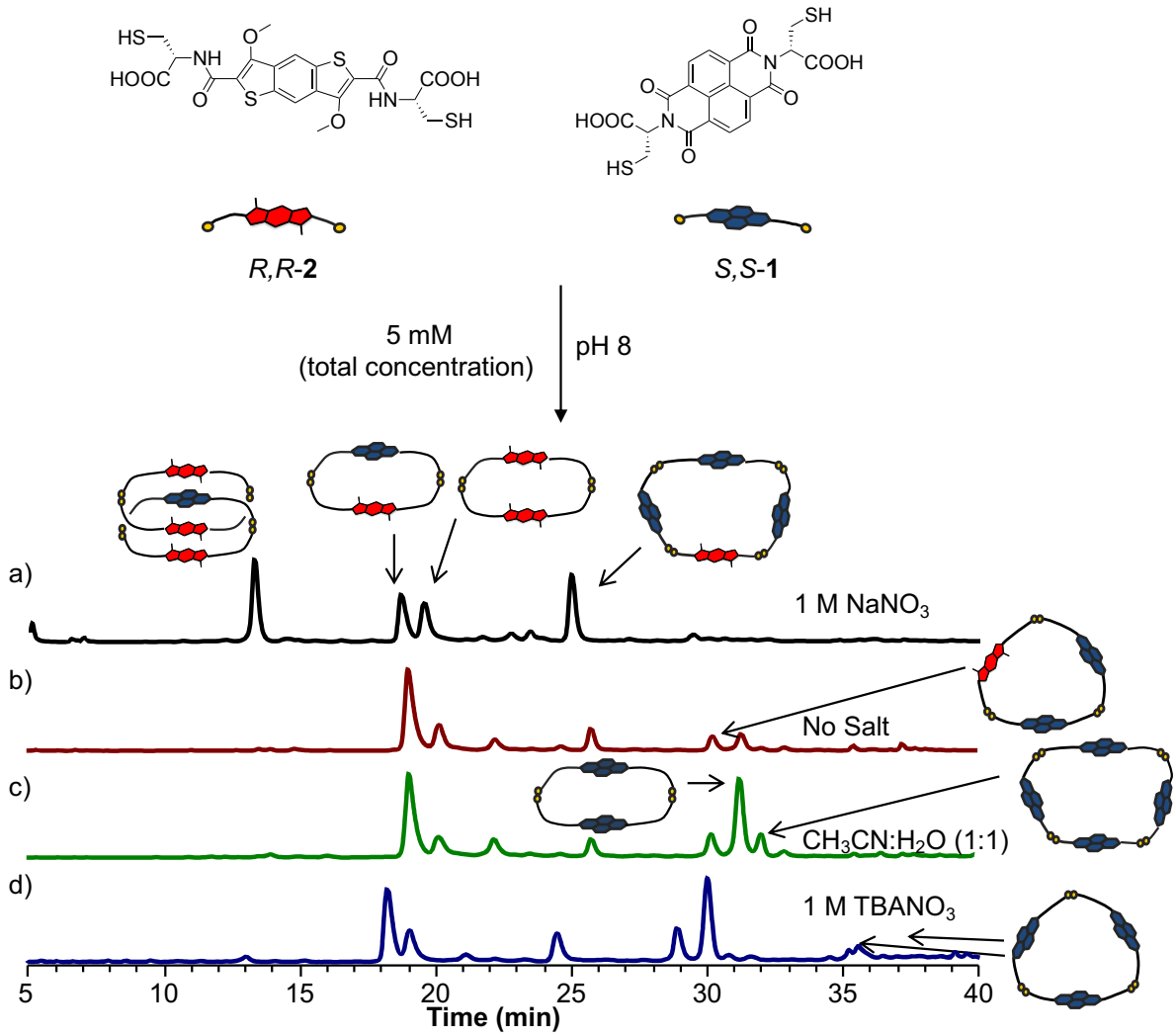


Figure 2.10. a) UV-Vis spectrum of **Cat I *RRRR*** at 23 °C (296 K), b) VT UV-Vis spectra of **Cat I *RRRR***, c) VT UV-Vis spectra of **Cat I *RRRR*** in CT region between 5 – 85 °C at the specified temperatures, d) CD spectrum of **Cat I *RRRR*** at 23 °C; e) VT CD spectra of **Cat I *RRRR*** and f) VT CD spectra of **Cat I *RRRR*** in CT region between 5 – 85 °C at the specified temperatures.

2.2. Analysis of *S,S*-1 and *R,R*-2 library

The building blocks, *S,S*-**1** and *R,R*-**2**, combined by themselves gave no catenated structure. One would expect that changing the chirality on NDI (from *R,R*-**1** to *S,S*-**1**), would have the same outcome in DCLs. When the library was repeated with 1 M $NaNO_3$ added, the presence of a peak at shorter retention time (than other species presented in the system) indicated the formation of a catenane (21%). By lowering the influence of the hydrophobic effect (setting the library in water : acetonitrile 1:1 mixture), or by introducing competing additives (1 M $TBANO_3$), only macrocycles were formed, as expected (see Figure 2.11 for the annotated chromatograms; see Figures 2.59 – 2.64 at the end of chapter for full characterization of each peak / component).



Species				
i.d.	**Cat II *RSRR***	**Y**	**Z**	***AAAD* (*SSSR*) macrocycle**
Species				
i.d.	***AAD* (*SSR*) macrocycle**	***S,S*-1 dimer**	***S,S*-1 trimer**	***S,S*-1 tetramer**

Figure 2.11. Reverse-phase HPLC analysis of *S,S*-**1**:*R,R*-**2** (1:1 molar ratio, 5 mM total concentration) library a) in the presence of 1 M $NaNO_3$, b) no salt, c) in a mixture of CH_3CN:H_2O (1:1) and d) in the presence of 1 M $TBANO_3$. Absorbances recorded at 389 nm. The unlabeled peaks did not ionise and could not be identified.

2.2.1. Analysis of the new identified *DADD* [2]catenane

The MS and MS/MS analyses (Figure 2.12) confirm that the molecule is a [2]catenane (**Cat II *RSRR***) formed by three *R,R*-**2** (BDT) units and one *S,S*-**1** (NDI) component – there are no species (in MS/MS) with *m/z* bigger than that of *R,R*-**2** homodimer.

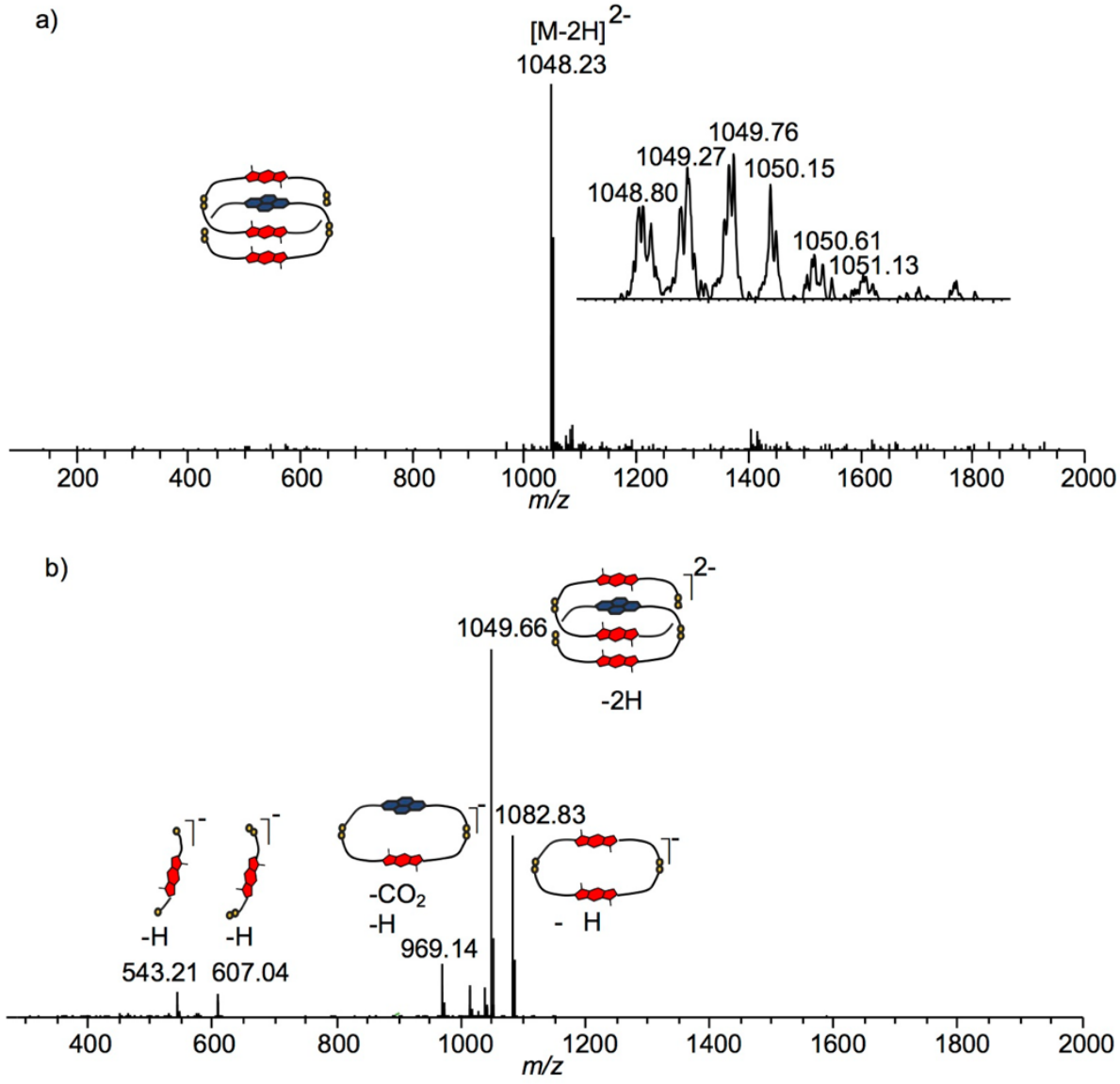
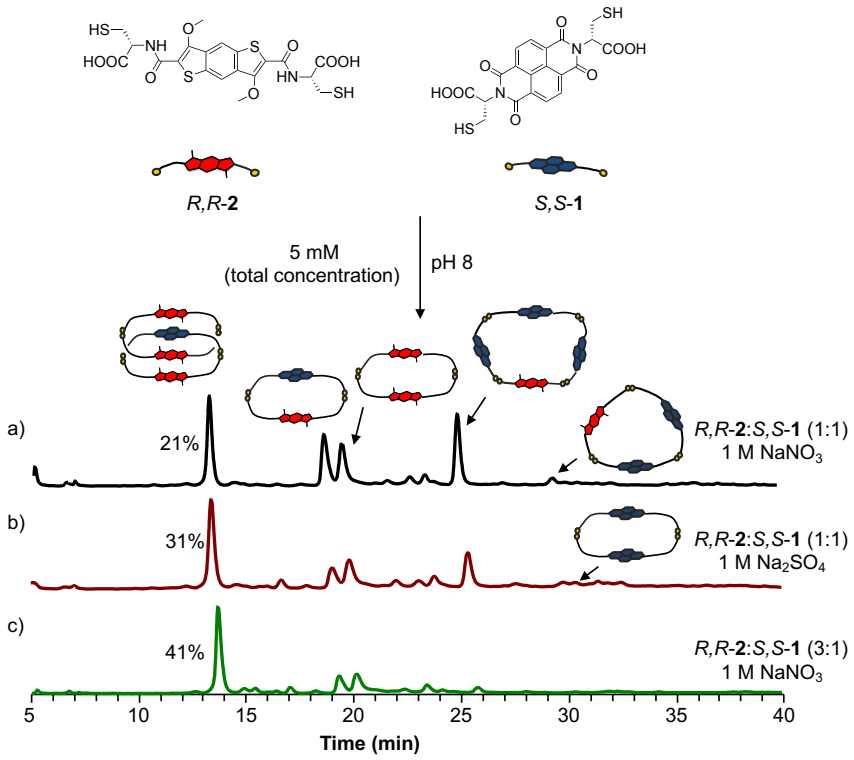


Figure 2.12. a) MS (*-ve*) of **Cat II *RSRR***; zoom of the molecular ion is shown as an inset and b) MS/MS (*-ve*) of **Cat II *RSRR***.

In an attempt to improve the yield of **Cat II *RSRR***, libraries were set with 1 M Na_2SO_4 (which is known to increase the hydrophobic effect more than $NaNO_3$, because of two sodium ions over one) and with building blocks in a 1:1 ratio

(*R,R*-**2**:*S,S*-**1**), which led to 31% yield. Another library with a 3:1 ratio of the same building blocks and 1 M $NaNO_3$ formed the catenane up to 41% (Figure 2.13).



Species				
i.d.	**Cat II *RSRR***	**Y**	**Z**	***AAAD* (*SSSR*) macrocycle**
Species				
i.d.	***AAD* (*SSR*) macrocycle**	*S,S*-**1 dimer**		

Figure 2.13. Reverse-phase HPLC analysis of *S,S*-**1**:*R,R*-**2** (1:1 molar ratio, 5 mM total concentration) library a) in the presence of 1 M $NaNO_3$ and b) 1 M Na_2SO_4 and c) *S,S*-**1**:*R,R*-**2** (1:3 molar ratio, 5 mM total concentration) in the presence of 1 M $NaNO_3$. Absorbance spectra recorded at 389 nm. The yields of **Cat II *RSRR*** are shown on top of each chromatogram. The unlabeled peaks did not ionise and could not be identified.

To further analyse the new catenane, it was isolated using semipreparative HPLC from a 3 mL library; the purity of the isolated sample is shown in Figure 2.14.

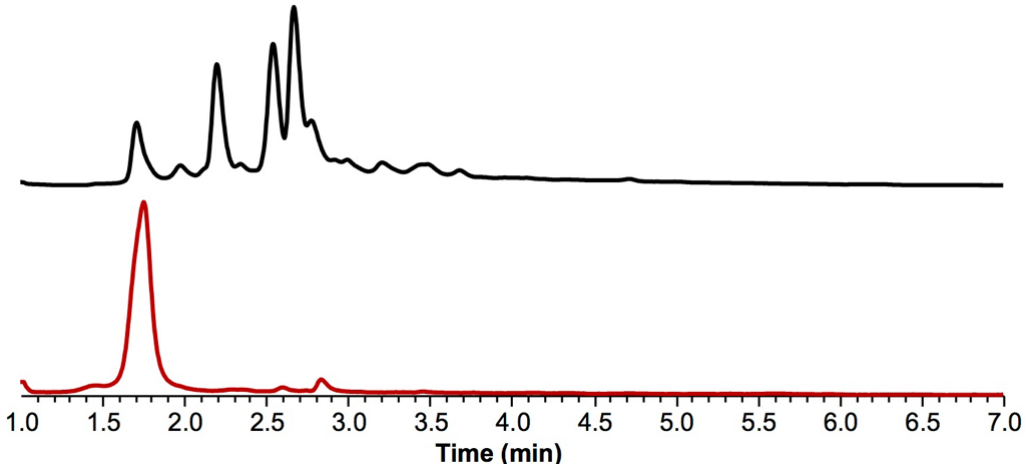


Figure 2.14. Reverse-phase HPLC analysis of *S,S*-**1**:*R,R*-**2** (1:1 molar ratio, 5 mM total concentration) library before (top trace) and **Cat II *RSRR*** after isolation (bottom trace). Absorbance spectra recorded at 389 nm.

The isolated sample was concentrated *in vacuo* and then analysed by ^{1}H NMR. The NMR shows three peaks corresponding to the BDT core at different chemical shifts; this is in line with the proposed structure. The most shielded peak is ascribed to the inner BDT proton (Figure 2.15, see below). With the use of NOESY (Figures 2.16 and 2.17), the remaining donor peaks were identified as being the peripheral BDT protons. The inner BDT proton (green) is correlated to the outer BDT proton (blue) and to NDI protons (black and magenta), which means that the NDI core is inside as well (also based on a lower chemical shift in comparison to that of the building block). The other BDT proton is considered to be outside the cavity (red). As in the previous case (**Cat I *RRRR***), the NH peaks are present and sharp, suggesting the quasi-pentacyclic structure. The VT-NMR experiment is useful in confirming that the structure is a catenane. In the case of an isomeric macrocycle, due to large thermal energy (??) at high temperature, the peaks in ^{1}H NMR spectrum will broaden. The other case shows that the peaks of the macrocycle broaden at room temperature, and, at high temperature, become sharp due to fast rotation and the signals average out. The sharpness of the peaks and the relative constant shift as the temperature increases is supportive of an interlocked geometry for this molecule. The HSQC spectrum (Figure

2.18) also suggests three different BDT protons (at different chemical shifts), but similar carbon chemical shifts.

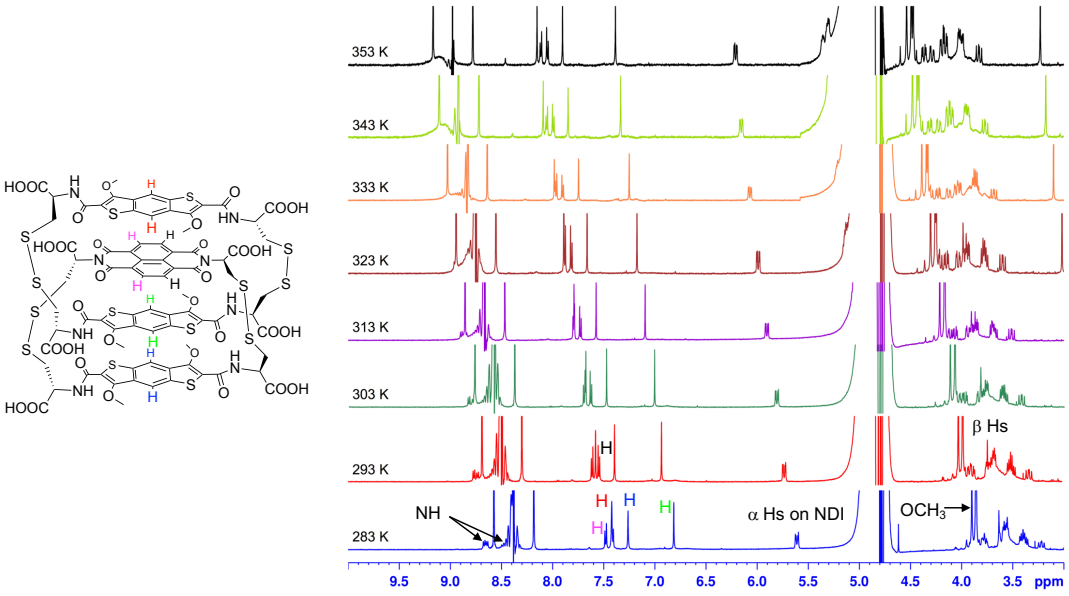


Figure 2.15. Variable temperature ^{1}H NMR (500 MHz) spectra of **Cat II *RSRR*** at the temperatures indicated above each spectrum. The solvent (H_2O) was referenced at 4.79 ppm. The spectrum was acquired in an H_2O:D_2O (~ 4:1) mixture.

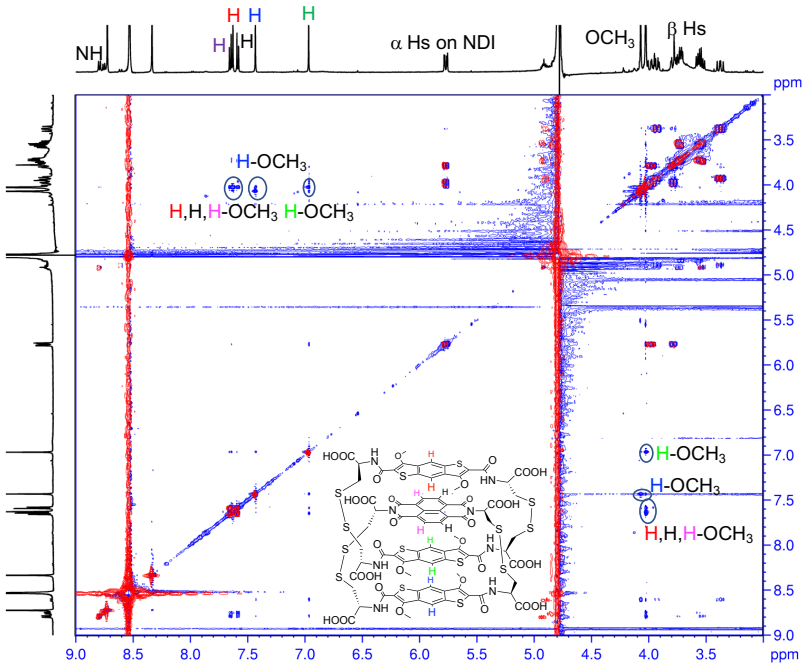


Figure 2.16. ^{1}H-^{1}H COSY (red) and 2D NOESY (blue) spectra superimposed of **Cat II *RSRR***. The solvent (H_2O) was referenced at 4.79 ppm. The spectrum was acquired in an H_2O:D_2O (~ 4:1) mixture.

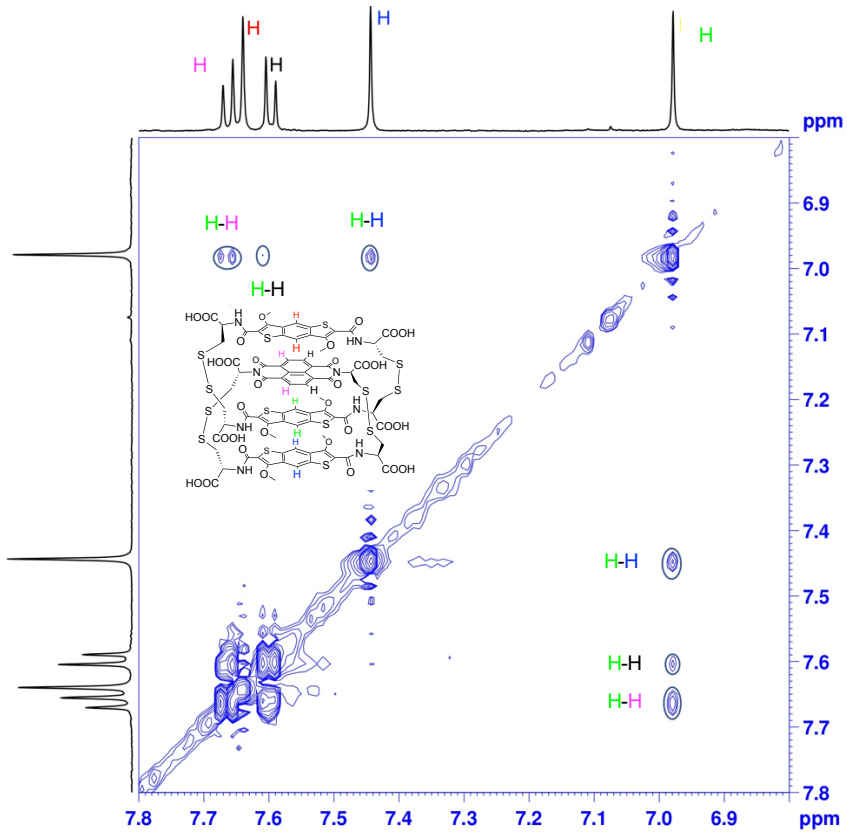


Figure 2.17. 2D NOESY spectrum of **Cat II *RSRR***. The solvent (H_2O) was referenced at 4.79 ppm. The spectrum was acquired in an H_2O:D_2O (~ 4:1) mixture.

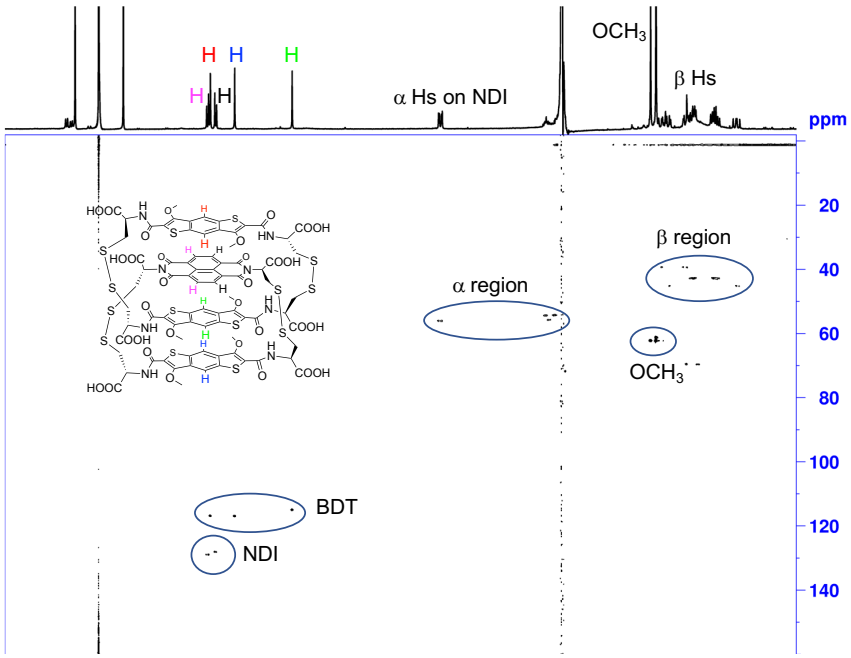


Figure 2.18. HSQC spectrum (500 MHz, 298 K) of **Cat II *RSRR***. The solvent (H_2O) was referenced at 4.79 ppm. The spectrum was acquired in an H_2O:D_2O (~ 4:1) mixture.

The presence of the CT band in both UV-Vis and CD spectra is a suggestion of the tightness / rigidity of the molecule. The broadness of the band (UV-Vis) indicates the orbital overlapping and also, by comparison, we notice that the chirality transfer in **Cat II *RSRR*** is less pronounced than in **Cat I *RRRR***. This may be because of an extra NDI in **Cat I *RRRR***, which is more rigid than the BDT. **Cat II *RSRR*** decomposes during the VT UV-Vis / CD experiments. As the temperature increases, the NDI / BDT bands start to shift, and the CD spectra show a clear transition with isosbestic points from product to the decomposed species. We suspect that the decomposition is caused by a combination of two factors: light and temperature (photo- and thermal degradation), as in the VT NMR experiment no decomposition is observed (Figure 2.19). There are also papers showing the decomposition of BDT derivatives at high temperatures (above 50 °C) or under light and exposure to air.[10–13]

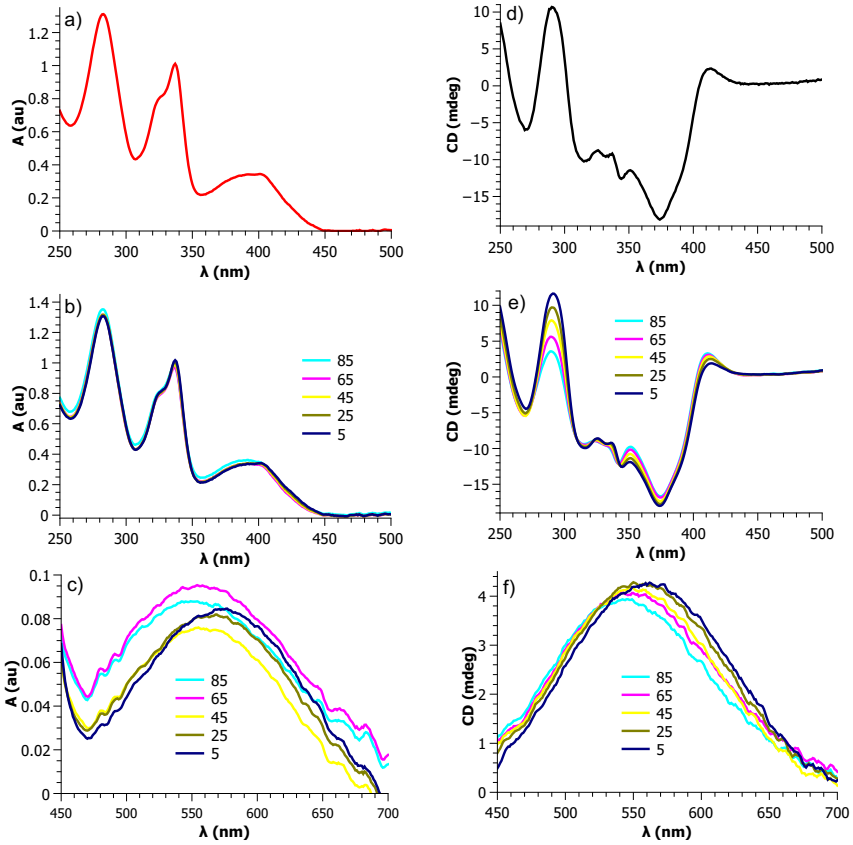


Figure 2.19. a) UV-Vis spectrum of **Cat II *RSRR*** at 23 °C, b) VT UV-Vis spectra of **Cat II *RSRR***, c) VT UV-Vis spectra of **Cat II *RSRR*** in the CT region between 5 – 85 °C at specified the temperatures, d) CD spectrum of **Cat II *RSRR*** at 23 °C; e) VT CD spectra of **Cat II *RSRR*** and f) VT CD spectra of **Cat II *RSRR*** in the CT region between 5 – 85 °C at the specified temperatures.

To complete the series of this study, a library with one racemic building block (NDI or BDT) and one enantiomer of the other (BDT or NDI, respectively) was set up.

2.3. Analysis of the libraries containing *R,R*-1, *S,S*-1 and *R,R*-2

A DCL containing a mixture of *R,R*-**1**, *S,S*-**1** and *R,R*-**2** (1:1:2 molar ratio) was set up under basic conditions and in the presence of 1 M $NaNO_3$, and stirred for three days in air-closed caped vials. The equilibrated library contains a large variety of species, including the [2]catenanes **Cat I *RRRR*** and **Cat II *RSRR*** as major components. The two [2]catenanes are non-isomeric and are produced in a structurally divergent process (Figure 2.20).

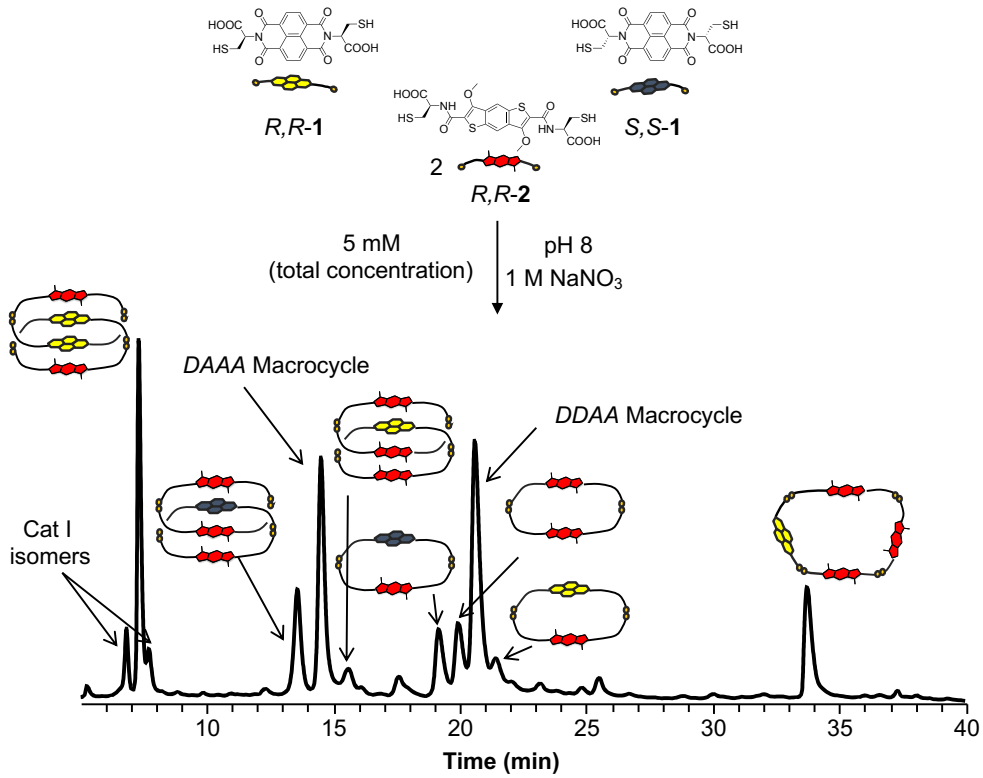
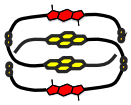
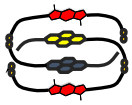
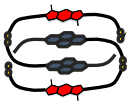
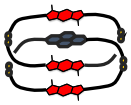
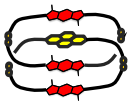
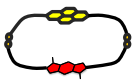
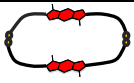


Figure is continued on the next page

Species					
i.d.	**Cat I *RRRR***	**Cat I *RRSR***	**Cat I *RSSR***	**Cat II *RSRR***	**Cat II *RRRR***
Amount in DCL	14%	5%		8%	2%
dr	75:14:11			77:23	
The total proportion of [2]catenanes: 29%					

Species			
i.d.	**X**	**Y**	**Z**
Amount in DCL	8%	9%	6%
The total proportion of dimers: 23%			

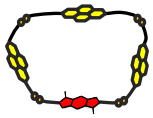
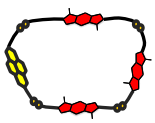
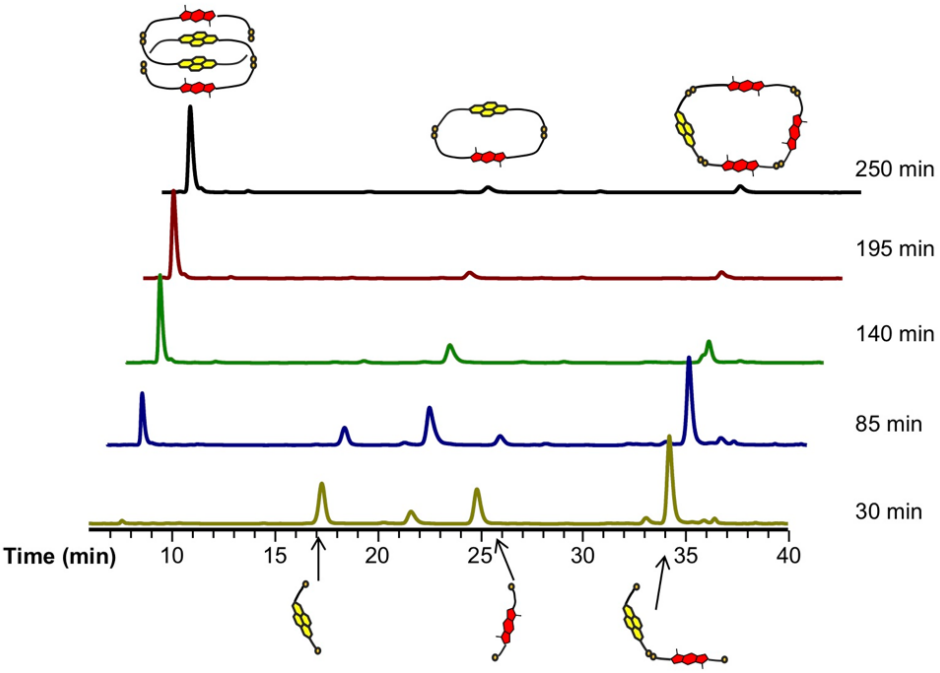
Species	*DAAA* I macrocycle	*DAAA* II macrocycle	*DDAA* macrocycle		
Amount in DCL	16%	2%	19%	2%	9%
The total proportion of tetramers: 48%					

Figure 2.20**.** Reverse-phase HPLC analysis at 389 nm of *R,R*-**1**:*S,S*-**1**:*R,R*-**2** (1:1:2 molar ratio, 5 mM total concentration) library in the presence of 1 M $NaNO_3$. The tables under the chromatogram show the yields of the individual species and the proportion of each class of macrocycles. For a complete description of the library, see Figure 2.65 (the nomenclature is based on Figure 2.25).

Now that we know how the 1:1:2 mixture library is behaving, we want to understand the reason behind this phenomenon. For this, we go back to the homochiral library (*R,R*-**1** with *R,R*-**2** in the presence of 1 M $NaNO_3$). The kinetic

evolution of this DCL (Figures 2.21 and 2.22) gives an outline of the mechanism beyond the formation of **Cat I *RRRR***. The first step is the formation of the linear heterodimer **X**, followed by its cyclisation. Then, the heterodimer **X** is threaded by the linear heterodimer, generating **Cat I *RRRR*** once the disulphide forms. Important note is that none of the homodimers are formed (either *R,R*-**1** dimer or *R,R*-**2** dimer – *i.e.* **Y**). On the basis of this observation and thermodynamics grounds, we can say that **X** is more stable than **Y**.



Cat I *RRRR*	**X**	***DDDA* (*RRRR*) macrocycle**	*R,R*-**1**	*R,R*-**2**	Linear **X**

Figure 2.21. Reverse-phase HPLC analysis of *R,R*-**1**:*R,R*-**2** (1:1 molar ratio, 5 mM total concentration) library over time in the presence of 1 M $NaNO_3$. Absorbances recorded at 389 nm at different intervals: 30 min, 85 min, 140 min, 195 min, 250 min, as shown in the figure.

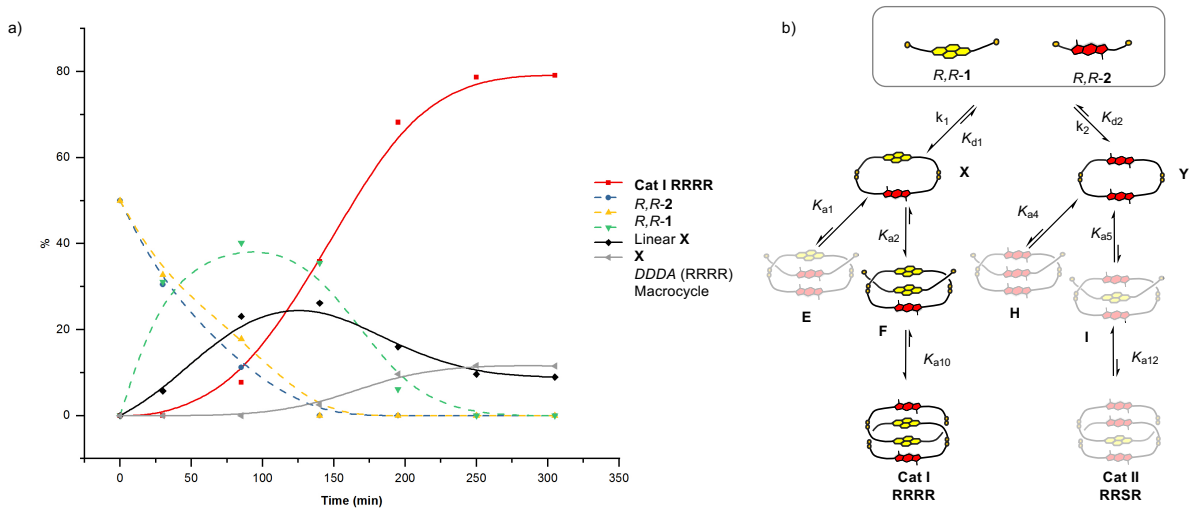
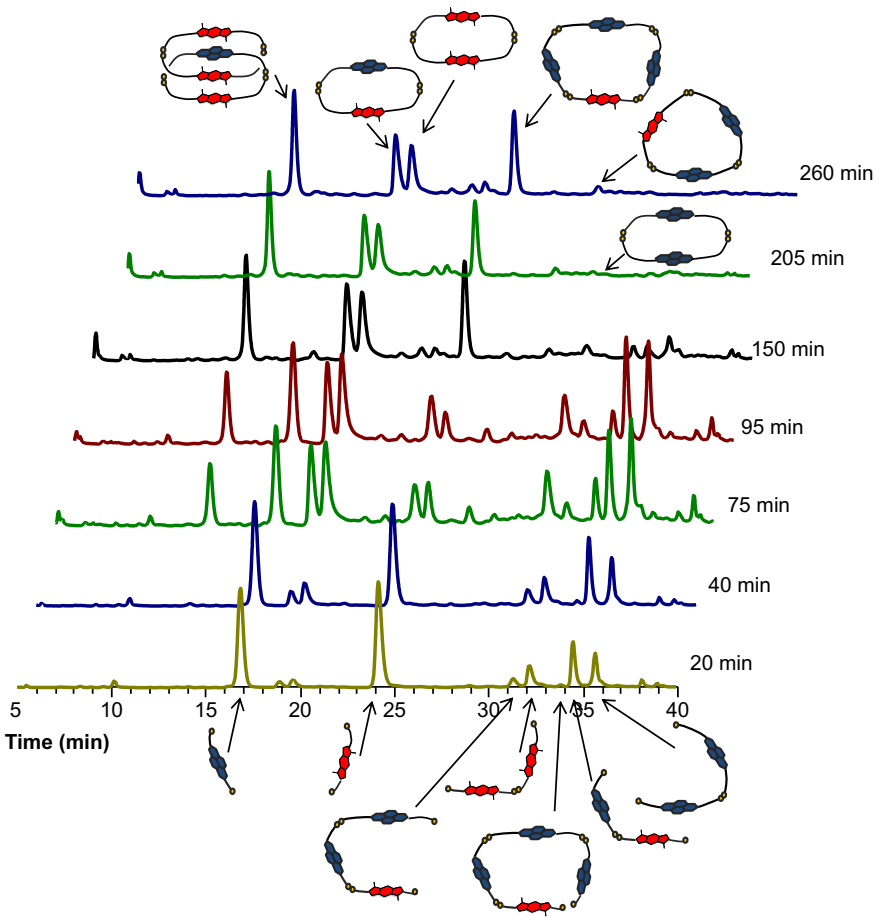


Figure 2.22. a) Kinetic profile of *R,R*-**1**:*R,R*-**2** (1:1, 5 mM total concentration, absorption monitored at 389 nm) showing the consumption of the building blocks (dashed blue, *R,R*-**2** and yellow-mustard lines, *R,R*-**1** respectively) and the formation of the **Cat I *RRRR*** (red line). The formation and disappearance of other species can also be tracked: linear **X** (dashed green), **X** (black) and ***DDDA* (RRRR) macrocycle** (grey). The dashed lines represent the intermediate species, while the solid lines the final products. *D* stands for donor (BDT) and *A* for acceptor (NDI) and b) the tree figure showing the pathway leading to **Cat I *RRRR***.

On the other side, the kinetic analysis (Figures 2.23 and 2.24) of library containing *S,S*-**1** and *R,R*-**2** (in the presence of 1 M $NaNO_3$) shows the formation of the corresponding heterochiral heterodimer **Z** and the homochiral homodimer **Y** (not seen in the previous homochiral library). Most likely, **Cat II *RSRR*** is formed by the threading of the dimer **Y** by the linear dimer **Z**. This process outcompetes the threading of the linear heterodimer in the closed dimer **Z** because the **Cat I *RSSR*** is not formed. Remarkably, in the homochiral library, **Cat II *RRRR*** is not formed, but instead, a macrocyclic tetramer is present (*DDDA* with *RRRR* stereochemistry).



Cat II *RSRR*	**Z**	**Y**	*S,S*-**1** dimer	***DAA* (*RSS*) macrocyle**	***DAAA* (*RSSS*) macrocycle**

S,S-**1**	*R,R*-**2**	Linear **Y**	Linear **Z**	Linear *S,S*-**1 dimer**	**Linear *DAA* (*RSS*)**	**Linear *DAAA* (*RSSS*)**

Figure 2.23. Reverse-phase HPLC analysis of *S,S*-**1**:*R,R*-**2** (1:1 molar ratio, 5 mM total concentration) library over time in the presence of 1 M $NaNO_3$. Absorbances recorded at 389 nm at different intervals: 20 min, 40 min, 75 min, 95 min, 150 min, 205 min, 260 min as shown in the figure.

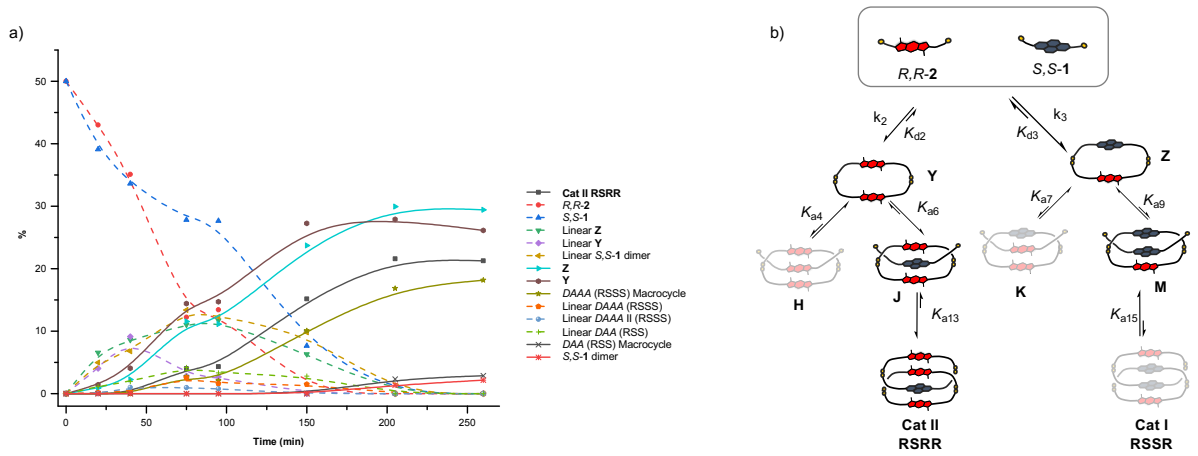


Figure 2.24. a) Kinetic profile of *S,S*-**1**:*R,R*-**2** (1:1, 5 mM total concentration, absorbance monitored at 389 nm) showing the formation of the **Cat II *RSRR*** (red). The description of each species can be seen in Legend (*D* stands for donor (BDT), and *A* stands for acceptor (NDI)) and b) the tree figure showing the pathway to **Cat II *RSRR***.

Once the kinetic and thermodynamic aspects of homo and heterolibraries are known, it is time to study the distribution of racemic DCLs over time. After 30 min, dimer **X** forms faster than either **Y** or **Z**, followed by its diastereomer **Y**. This is further amplified, and after 85 min, **X** represents 13% of the DCL, while **Z** is present only in 4%. On the other hand, the homochiral homodimer **Y** constitutes 8% of the library. Based on this, the structural divergence is due to two factors: the higher stability of two (**F** and **J**) out of the nine possible pseudorotaxanes (Figure 2.25), and the relative stability of the final products (**Cat I *RRRR*** and **Cat II *RSRR***).

The first step in the catenation process is the formation of the linear dimers followed by their cyclisation. In the *R,R*-**1**:*S,S*-**1**:*R,R*-**2** 1:1:2 library, **X** forms in larger proportion than **Z**, suggesting it is more thermodynamically stable. This means that changing the chirality leads to a less pronounced aromatic stacking, hence lower stability of dimer **Z**. In the homochiral library, **Y** is not present at all, whereas in the heterochiral one, it is the main final product. This implies that **Z** is the least stable between all three dimers (**Z** is less stable than **Y** because, based on the statistical ground, the ratio **Y**:**Z** should be 1:2, which is not valid in this case). Also, in Figure 2.25, the NDI homodimers are not present, which is consistent with previous studies, which showed that their cavity is too small to allow threading.[9]

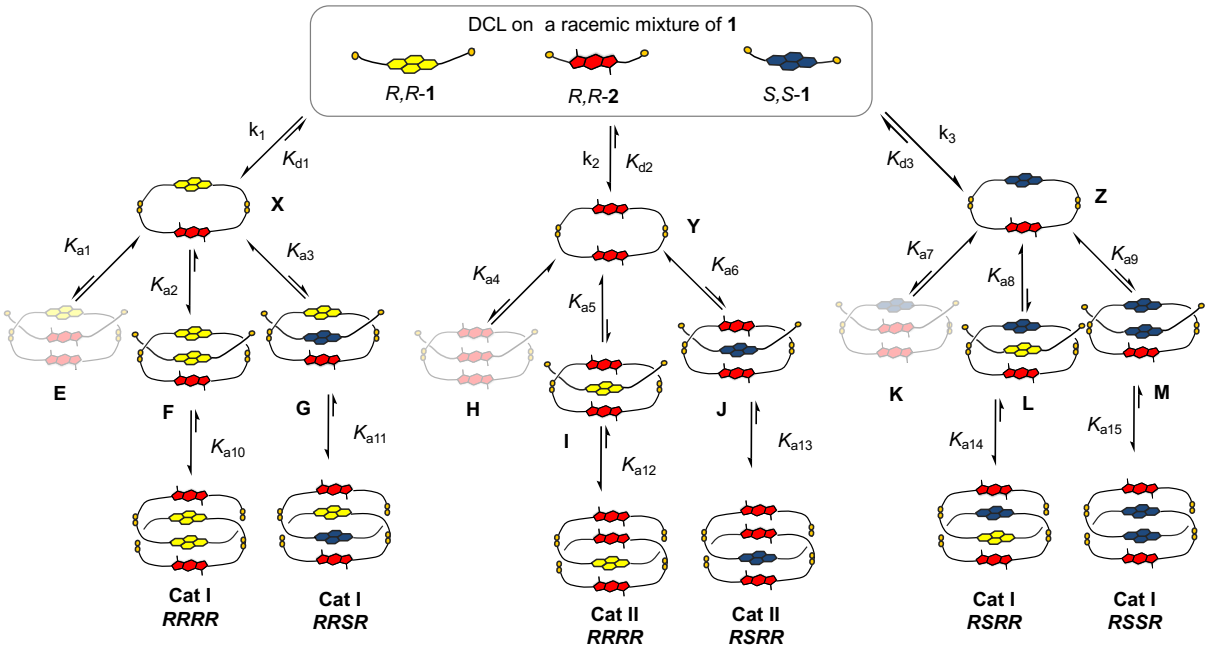
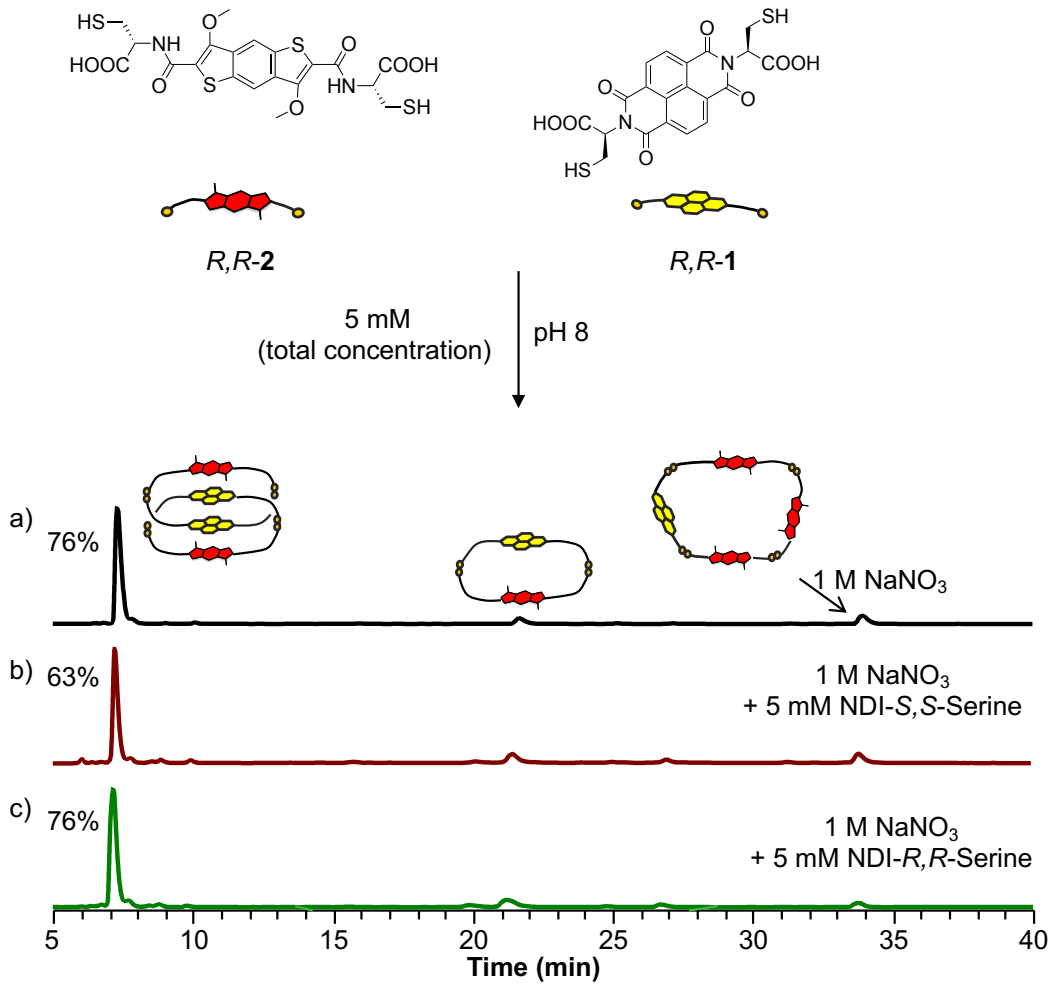


Figure 2.25. Schematic representation (simplified) of the pathways leading to **Cat I *RRRR*** and **Cat II *RSRR***. The equilibrium arrows show the relative position of the equilibrium based on theoretical and experimental data. The faded structures are either present in less than 1% or not present in the DCLs analysed. K_d = dimerization constant, K_a = association constant.

The next step is the threading of an NDI or BDT unit through the cavity of the dimers. The optimum donor-acceptor interaction determines the best stacking. Thus, **X** is a better threading partner for NDI than for BDT. Between *R,R*-**1** and *S,S*-**1**, the first one is more favourable for threading based on chirality grounds and on the formation of **Cat I *RRRR*** over its diastereomers. This observation was strengthened by control experiments, in which competitors were used in the libraries (NDI-*S,S*-serine and NDI-*R,R*-serine, which are analogues of *R,R*-**1** and *S,S*-**1**, respectively). For the DCL containing *R,R*-**1** and *R,R*-**2** (in the presence of 1 M $NaNO_3$), NDI-*S,S*-serine lowers the yield of **Cat I *RRRR*** in comparison to NDI-*R,R*-serine (Figure 2.26).

DCLSim is a software developed at the University of Cambridge, which can simulate DCL distributions based on an initial set of parameters such as energies of formation and association constants with a template. DCLSim[14] simulations agreed with the experimental data (see the tables at the end of the Chapter). In conclusion,

the following trend applies to the left-hand side section of the Figure 2.25: $K_{a2} > K_{a3} >> K_{a1}$.



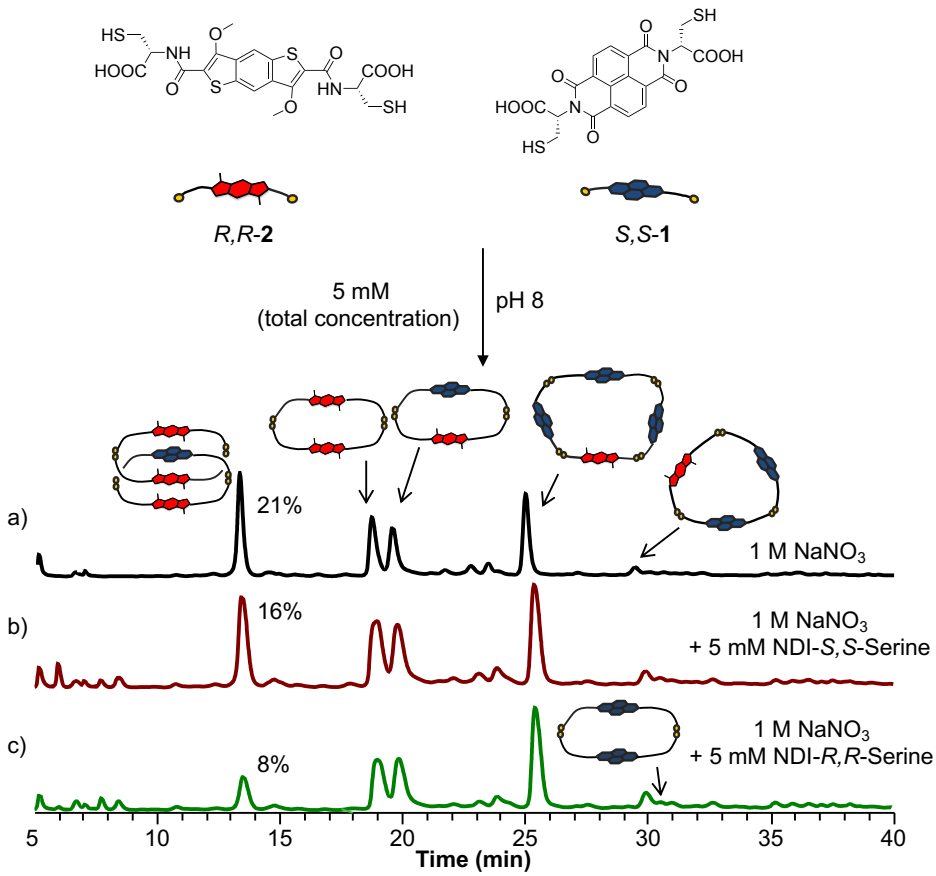
Species			
i.d.	**Cat I *RRRR***	**X**	***AAAD* (*RRRR*) macrocycle**

Figure 2.26. Reverse-phase HPLC analysis of *R,R*-**1**:*R,R*-**2** (1:1 molar ratio, 5 mM total concentration) library a) in the presence of 1 M $NaNO_3$, b) 1 M $NaNO_3$ + 5 mM NDI-*S,S*-serine and c) 1 M $NaNO_3$ + 5 mM NDI-*R,R*-serine. Absorbance recorded at 389 nm. The unlabelled peaks did not ionise and could not be identified.

In the case of the heterochiral DCL, the most stable pseudorotaxane is **J** based on donor-acceptor interaction. On the same basis, **J** is still the most stable pseudorotaxane (between **H**, **I** and **J**) in the *R,R*-**1**:*S,S*-**1**:*R,R*-**2** 1:1:2 library, based on

the amount of **Cat II *RRRR*** formed (~2%). The control experiments (using NDI-*S,S*-serine and NDI-*R,R*-serine) support the statements that NDI-*R,R*-serine lowers the yield of **Cat II *RSRR*** more than its enantiomer (Figure 2.27). Moreover, UV-vis titrations of **Y** (the single major component in the DCL of *R,R*-**2**) with the competitors also show a higher association constant for pseudorotaxane **J** formation than for **I** (Figure 2.28). For the titrations, the competitors were used instead of building blocks as these will not regenerate the DCL. The use of building blocks will reinitiate the DCL – this shows that the DCL is under thermodynamic control and is not a kinetic trap – Figures 2.29 – 2.32. Based on these experiments, we can conclude that $K_{a6} > K_{a5} > K_{a4}$. **H** has the lowest association constant as it consists of three stacked donor units, which is energetically unfavourable. The DCLSim simulations also agree with the results obtained (See Tables at the end of the chapter), which indicate that **M** is more stable than **L** ($K_{a9} > K_{a8}$). Also, it is worth mentioning that **Cat II *RRRR*** cannot be formed through pathway **E** as **X** is consumed for the formation of **Cat I *RRRR***, and also **Y** is not formed in the homochiral DCL.

As for the other pseudorotaxanes **E**, **K** and **H**, they are less likely to form as they have donor-donor interactions, which are less favourable than acceptor-acceptor interactions. **H,** formed by three donors stacked together, is less likely to form therefore K_{a1}, $K_{a7} > K_{a4}$. In an attempt to increase the catenane formation, DCLs with building blocks in different ratios were set up (Figure 2.33). The following ratios between *R,R*-**2**, *R,R*-**1** and *S,S*-**1** were used: 5:1:1, 1:1:1, 1:2:2 and 5:2:1. Unfortunately, varying the ratio of building blocks does not improve the yield for **Cat I *RRRR*** and **Cat II *RSRR*** formation, but rather leads to a general increase in yield of all catenated species in the system. This has even had an opposite effect, in some cases the catenanes of interest not being formed at all. In the first case (Figure 2.33a), the percentages of all catenanes increase due to the excess of *R,R*-**2**. In the second and third examples, **Cat II** is missing since *R,R*-**2** is consumed entirely in the formation of other species. It is important to note that the last case (*R,R*-**2**:*R,R*-**1**:*S,S*-**1** 5:2:1) is not a racemic mixture (*R,R*-**1**:*S,S*-**1** ratio is 2:1). **Cat II *RRRR*** is increased since *R,R*-**1** is in excess with respect to its enantiomer.



Species				
i.d.	**Cat II *RSRR***	**Y**	**Z**	***AAAD* (*SSSR*) macrocycle**
Species				
i.d.	***AAD* (*SSR*) macrocycle**	***S,S*-1 dimer**		

Figure 2.27. Reverse-phase HPLC analysis of *S,S*-**1**:*R,R*-**2** (1:1 molar ratio, 5 mM total concentration) library a) in the presence of 1 M $NaNO_3$, b) 1 M $NaNO_3$ + 5 mM **NDI-*S,S*-serine** and c) 1 M $NaNO_3$ + 5 mM **NDI-*R,R*-serine**. Absorbances recorded at 389 nm. The unlabelled peaks did not ionise and could not be identified.

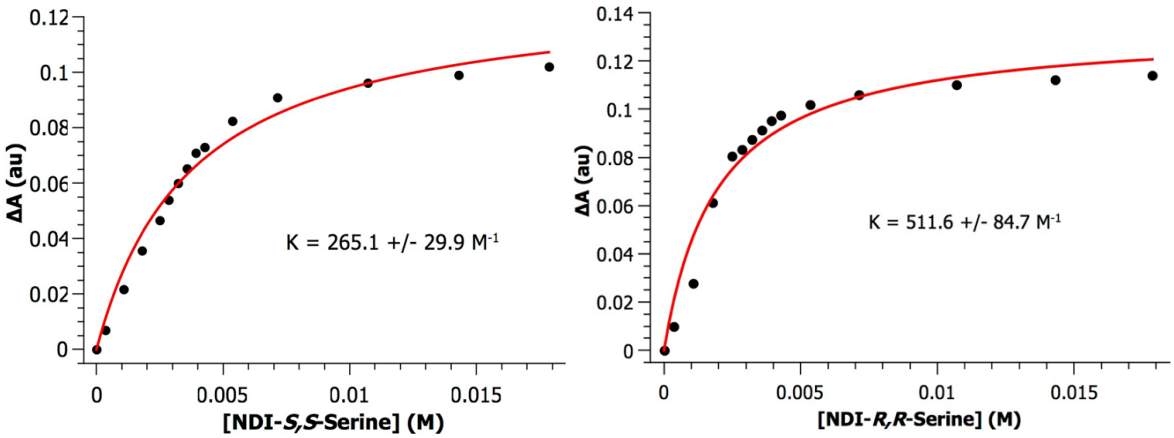
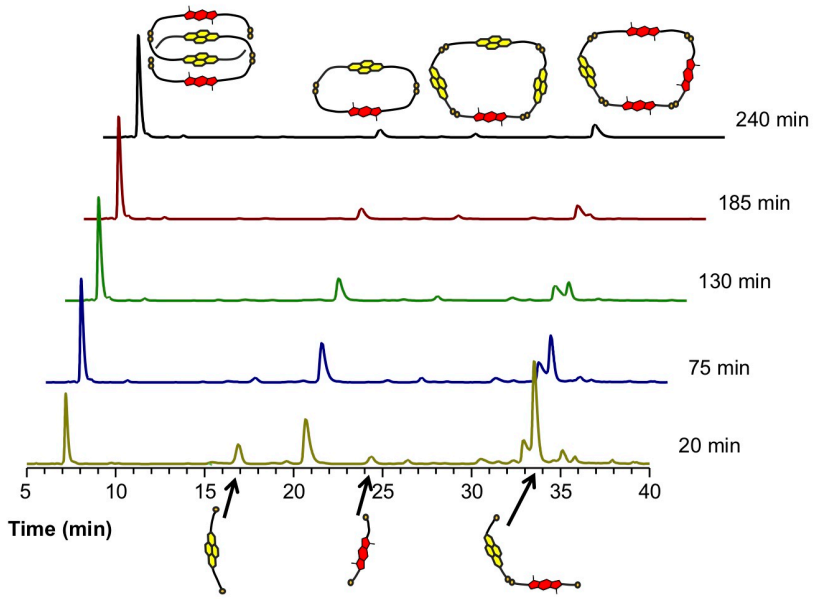


Figure 2.28. Fitting of UV-vis titration data points of the binding event between **Y** and **NDI-*S,S*-serine** (left). Fitting of UV-vis titration data points of the binding event between **Y** and **NDI-*R,R*-serine** (right). Experiment were done in water in the presence of 1 M $NaNO_3$ at 23 °C.



Cat I *RRRR*	**X**	***DAAA* (*RRRR*) macrocycle**	***DDDA* (*RRRR*) macrocycle**	*R,R*-**1**	*R,R*-**2**	Linear **X**

Figure 2.29. Kinetic profile of *R,R*-**1**:**Y** over time in the presence of 1 M $NaNO_3$. Absorbances recorded at 389 nm at different intervals: 20 min, 75 min, 130 min, 185 min, 240 min as shown in figure.

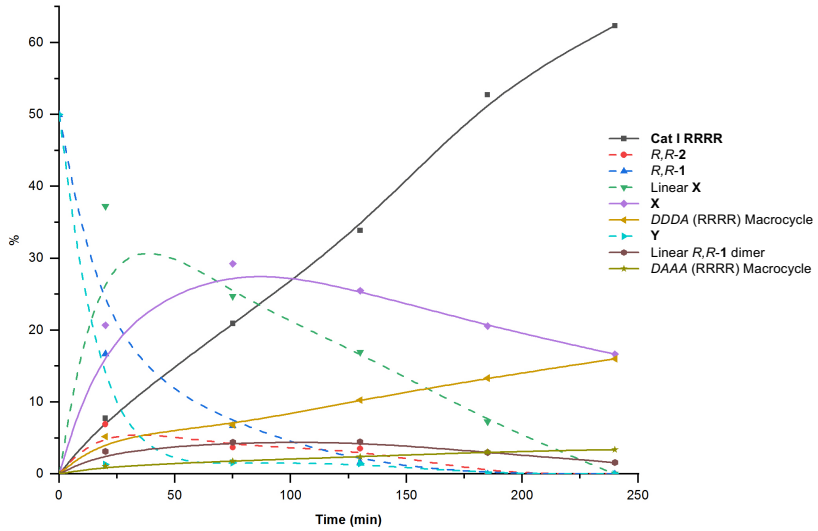


Figure 2.30. Kinetic profile of *R,R*-**1**:**Y** (absorbance monitored at 389 nm) showing the formation of the **Cat I *RRRR*** (black line). The description of each species is given in the legend (*D* stands for donor (BDT) and *A* stands for acceptor (NDI)).

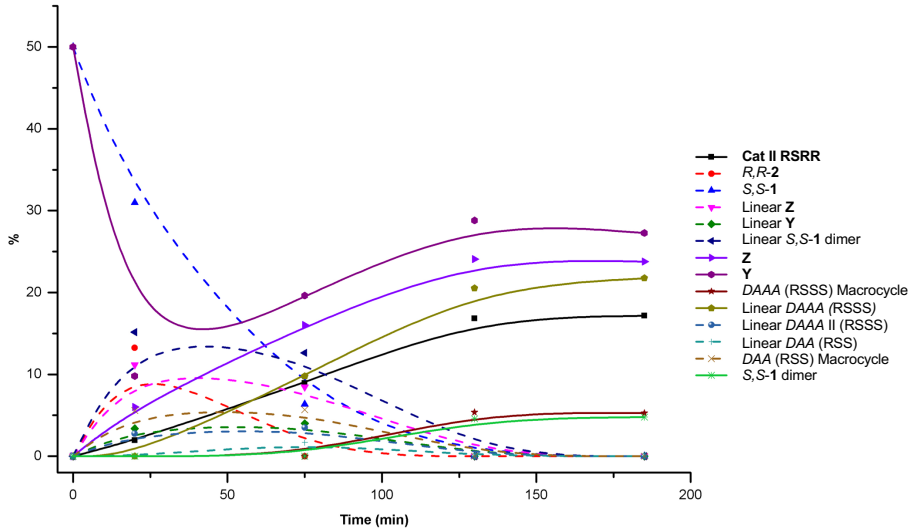
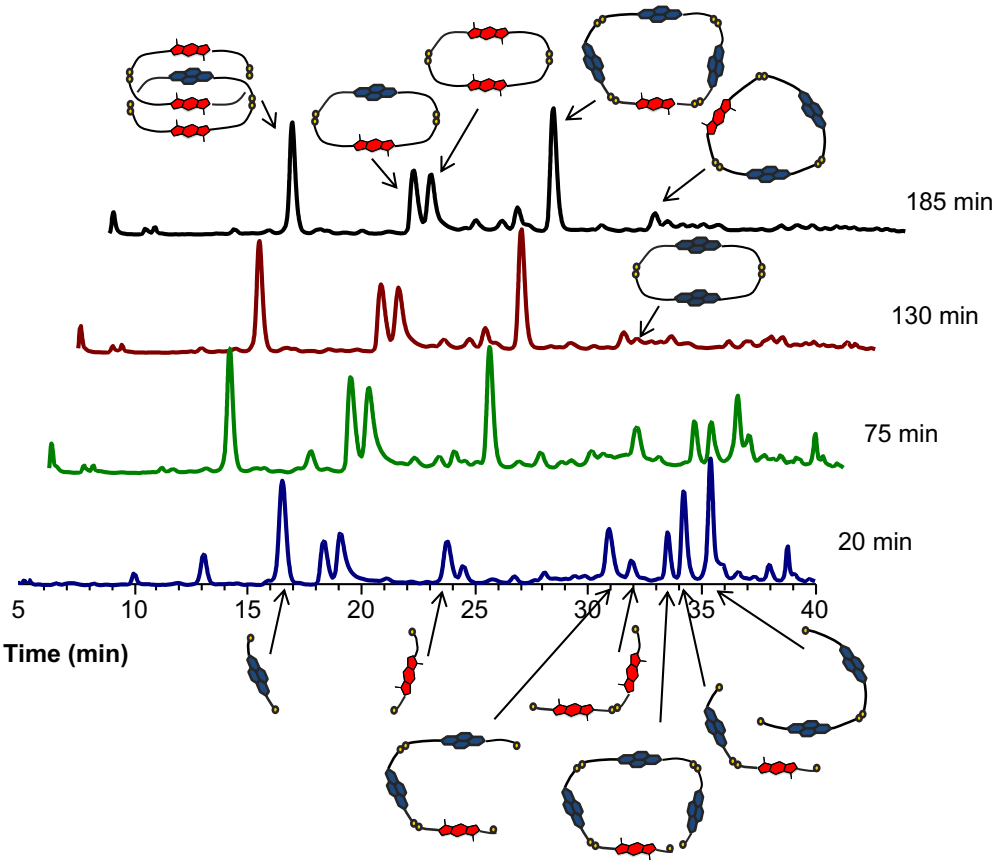


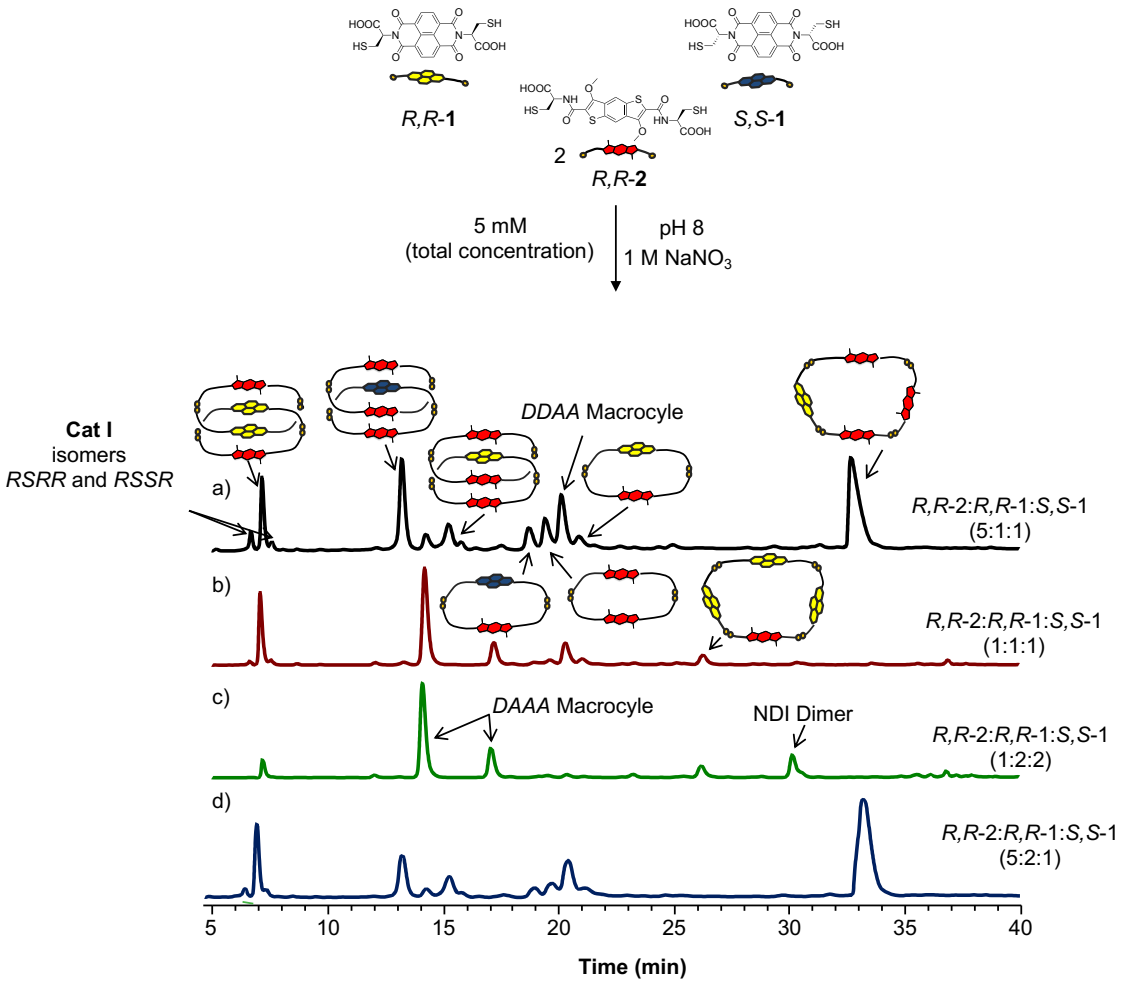
Figure 2.31. Kinetic profile of *S,S*-**1**:**Y** 1:1 (absorbance monitored at 389 nm) showing the formation of the **Cat II *RSRR*** (black line). The description of each species is given in the legend (*D* stands for donor (BDT) and *A* stands for acceptor (NDI)). Figure added here for space requirements. This is the kinetic study graphic representation for the next figure.



Cat II *RSRR*	**Z**	**Y**	*S,S*-**1** dimer	***DAA* (*RSS*) macrocyle**	***DAAA* (*RSSS*) macrocycle**

S,S-**1**	*R,R*-**2**	Linear **Y**	Linear **Z**	Linear *S,S*-**1 dimer**	**Linear *DAA* (*RSS*)**	**Linear *DAAA* (*RSSS*)**

Figure 2.32. Kinetic profile of *S,S*-**1**:**Y** 1:1 over time in the presence of 1 M $NaNO_3$. Absorbances recorded at 389 nm at different intervals: 20 min, 75 min, 130 min, 185 min as shown in the figure.



Species					
i.d.	**Cat I *RRRR***	**Cat I *RRSR***	**Cat I *RSSR***	**Cat II *RSRR***	**Cat II *RRRR***

Species			
i.d.	**X**	**Y**	**Z**

Figure is continued on the next page

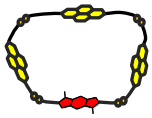
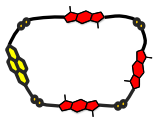
Species		
i.d.	***AAAD* (*RRRR*) macrocycle**	***DDDA* (*RRRR*) macrocycle**

Figure 2.33. Reverse-phase HPLC analysis of a) *R,R*-**2**:*R,R*-**1**:*S,S*-**1** (5:1:1 molar ratio, 5 mM total concentration) library in the presence of 1 M $NaNO_3$, b) *R,R*-**2**:*R,R*-**1**:*S,S*-**1** (1:1:1 molar ratio, 5 mM total concentration) library in the presence of 1 M $NaNO_3$, c) *R,R*-**2**:*R,R*-**1**:*S,S*-**1** (1:2:2 molar ratio, 5 mM total concentration) library in the presence of 1 M $NaNO_3$ and d) *R,R*-**2**:*R,R*-**1**:*S,S*-**1** (5:2:1 molar ratio, 5 mM total concentration) library in the presence of 1 M $NaNO_3$. Absorbances recorded at 389 nm. The unlabelled peaks did not ionise and could not be identified.

The last part of the study was to calculate the stability of final components through computational calculations, as this is one of the factors that determine the final distribution in a DCL. The calculations (PM7, COSMO water model) showed that **Cat I *RRRR*** and **Cat II *RSRR*** are the most stable components among their diastereomers. **Cat I *RRRR*** is more stable than **Cat I *RSSR*** with about 36.5 kJ mol^{-1} and with about 50 kJ mol^{-1} than **Cat I *RSRR***. In the other case, **Cat II *RSRR*** is more stable than its diastereomer (**Cat II *RRRR***) with about 9 kJ mol^{-1}. The formation of **Cat II *RSRR*** is more favourable because of the preferential threading of *R,R*-**1** in **X**. This leads to the formation of **Cat I *RRRR*** at the expense of **Cat II *RRRR*** in the *R,R*-**1**:*S,S*-**1**:*R,R*-**2** 1:1:2 DCL. The stacked structure imposes the carboxylate moieties to be in close proximities, so the overall structure will have new steric and electrostatic repulsion when the chiral centre is changed (Table 2.1).

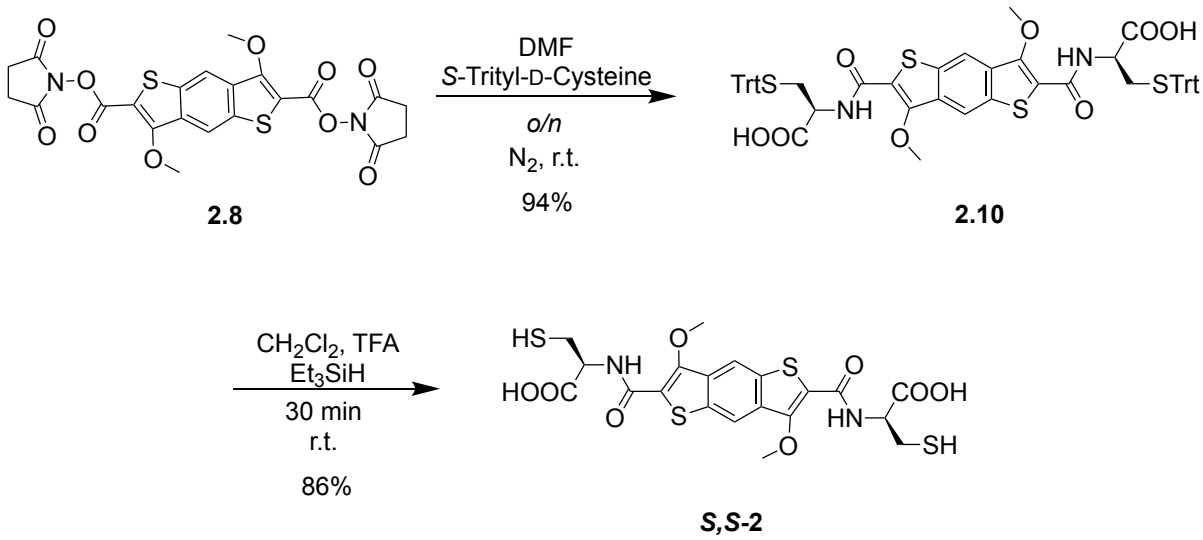
Table 2.1. Energies of different types of **Cat I** and **Cat II**.

Catenane	Energy (kJ mol^{-1})
DAAD **Cat I *RRRR***	-8530.72
DAAD **Cat I *RSSR***	-8494.23
DAAD **Cat I *RSRR***	-8480.76
DADD **Cat II *RRRR***	-8501.01
DADD **Cat II *RSRR***	-8509.83

As a mini-conclusion for this part, the first ever structural divergence reaction on racemic mixture using DCC is reported herein. The reaction takes place in water, at room temperature and does not require expensive metal catalysts. This is due to the stereospecific formation of pseudorotaxane and thermodynamic stability of final structures and intermediate species.

The first part of this chapter described the behaviour of the homochiral and heterochiral libraries containing *R,R*-**2** and **1**. We have also discussed the outcome of DCLs containing *R,R*-**2** and a racemic mixture of NDI (*S,S*-**1** and *R,R*-**1**). In the following part, we will discuss the enantiomeric libraries containing *S,S*-**2** and **1**, followed by the analysis of DCLs containing *R,R*-**1** and a racemic mixture of **2**.

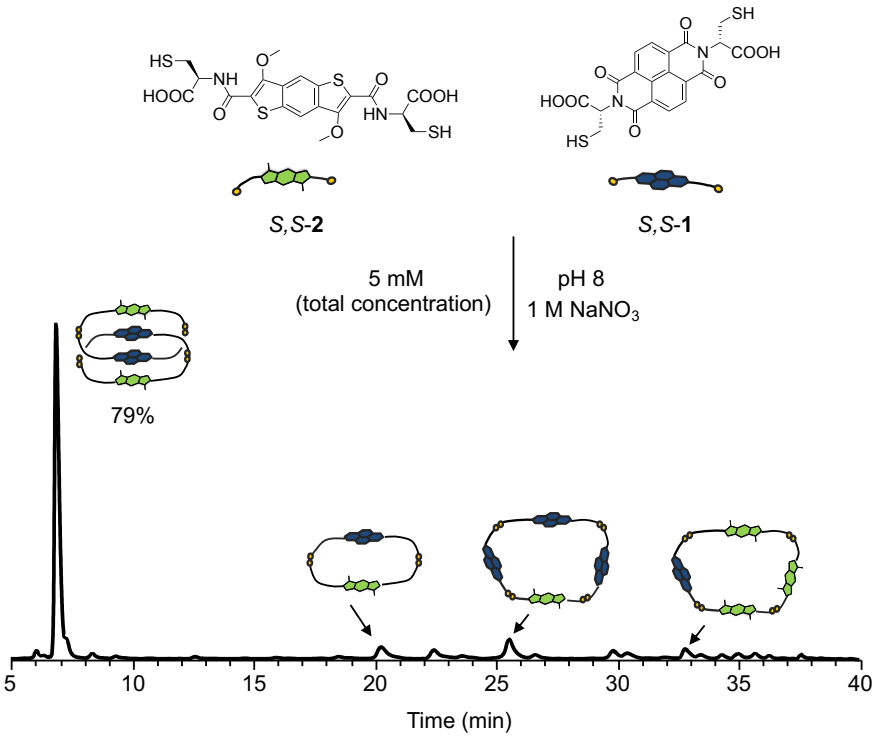
The synthetic route for making *S,S*-**2** is similar to *R,R*-**2**. It starts from compound **2.8** shown in Scheme 2.3, which is reacted with *S*-trityl-D-cysteine, followed by the cleavage of the trityl group to give the free thiol molecule *S,S*-**2** in similar yield to *R,R*-**2** (Scheme 2.3). The Q_{zz} value of this building block is the same as for its enantiomer as no new moieties are attached to the aromatic core.



Scheme 2.3. The synthetic scheme of *S,S*-**2**.

2.4. Analysis of *S,S*-1 and *S,S*-2 library

Our studies on the other enantiomer have showed that the structural divergence is accentuated at high concentrations of salt (the hydrophobic effect is increased). On this basis, the new libraries will be done only in the presence of 1 M $NaNO_3$. It is expected the homo and heterochiral DCLs should behave similarly to the previous ones because the building blocks are enantiomers (in relation to the systems described earlier in this chapter). As expected, the homochiral DCL (*S,S*-**1** and *S,S*-**2**) yields the **Cat I *SSSS*** in a similar yield (79%). The only small difference is that, besides a *DDDA* macrocycle, a *DAAA* macrocycle forms. The mechanism is likely to follow the one described for *R,R*-**1** and *R,R*-**2**, as this library is enantiomeric. The cyclic homochiral heterodimer **X'** is formed firstly, then its linear version is threaded through. The last step is the disulphide formation to give **Cat I *SSSS*** (Figure 2.34).

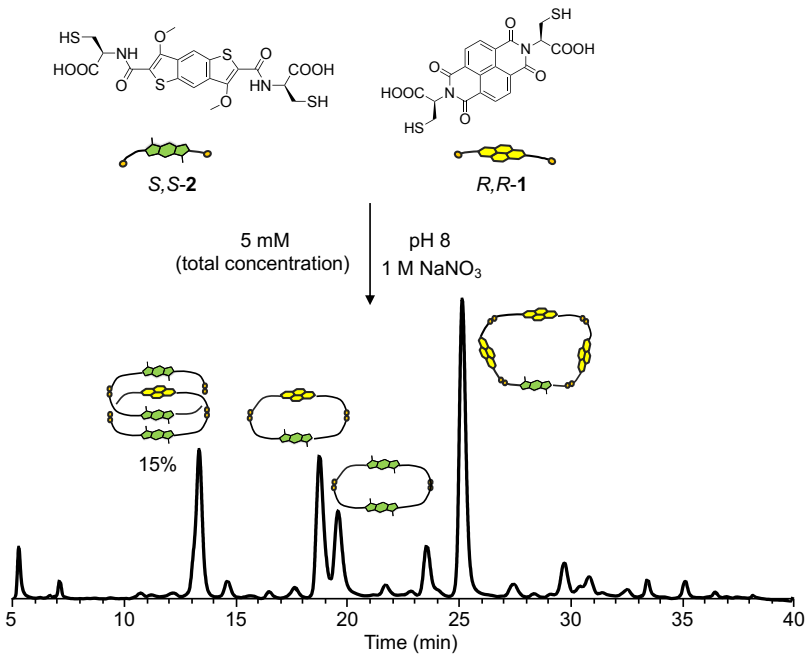


Species				
i.d.	**Cat I *SSSS***	**Z1**	***AAAD* (*SSSS*) macrocycle**	***DDDA* (*SSSS*) macrocycle**

Figure 2.34. Reverse-phase HPLC analysis of *S,S*-**2**:*S,S*-**1** (1:1 molar ratio, 5 mM total concentration) library in the presence of 1 M $NaNO_3$.

2.5. Analysis of *R,R*-1 and *S,S*-2 library

The distribution of the heterochiral library (*R,R*-**1** and *S,S*-**2**) shows four main products with corresponding yields given in brackets: **Cat II *SRSS*** (15%), homodimer **Y2** (30%), heterochiral heterodimer **Z'** (20%) and *DAAA* tetramer (28%). Based on the fact that the library is enantiomeric to *S,S*-**1** and *R,R*-**2** one, the mechanisms of this libraries are likely to be similar. Firstly, **Y2** is formed followed by the threading of linear *R,R*-**1**. The last step is to close up the ring to form **Cat II *SRSS*** (Figure 2.35).



Species				
i.d.	**Cat II *SRSS***	**X1**	**Y2**	***AAAD* (*SSSS*) macrocycle**

Figure 2.35. Reverse-phase HPLC analysis of *S,S*-**2**:*R,R*-**1** (1:1 molar ratio, 5 mM total concentration) library in the presence of 1 M $NaNO_3$.

2.6. Analysis of the libraries containing *R,R*-1, *S,S*-2 and *R,R*-2

The outcome of the library containing *R,R*-**1** and a racemic mixture of **2** in a 2:1:1 ratio (Figure 2.36) is complex and significantly different from the diastereomeric DCLs discussed above (pages 67 – 68). The initial analysis shows the formation of two types of **Cat I**: **Cat I *RRRR*** formed in the largest proportion (42%), and its diastereomer **Cat I *RRRS*** (12%). The stereochemistry for the latter was assigned as *R,R,R,S* due to the fact that as it was shown previously (page 26) that the heterochiral heterodimer is less likely to form to a [2]catenane due to chirality mismatch when compared to the homochiral heterodimer (see Figure 2.37).

The diastereomers of **Cat II** are the next most abundant species present in this DCL. Based on the LC-MS analysis, seven of the eight possible stereoisomers of

Cat II are present in the mixture (see Figure 2.38 for a diagram showing all the pathways going to **Cat II** isomers). This is in striking contrast with the diastereomeric DCL containing ***R,R*-2** and a racemic mixture on **1**, where the **Cat II** ***RSRR*** was the dominant species in a mixture of two separable diastereomers. This remarkable outcome further adds to the complexity unravelled by SDRRM DCLs.

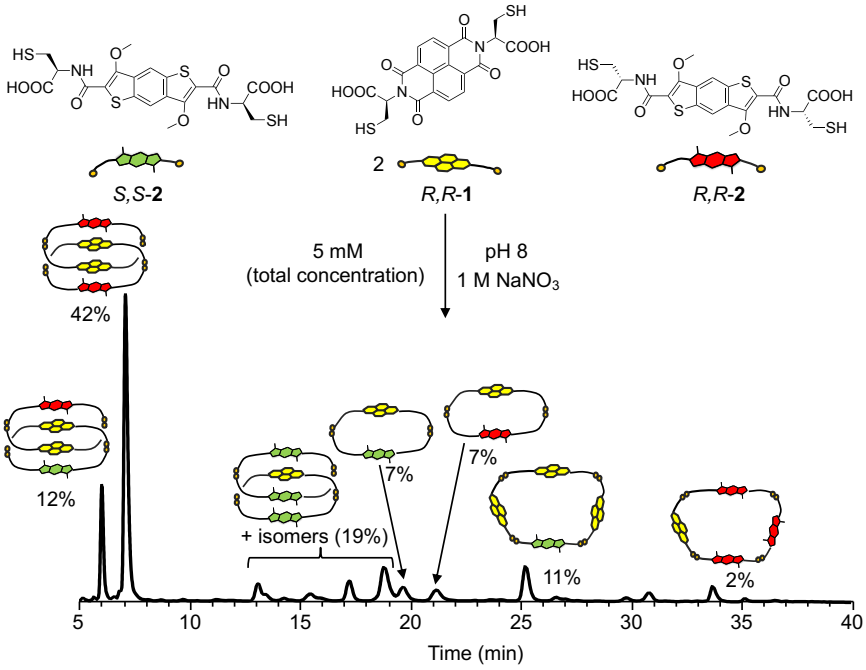


Figure 2.36. Reverse-phase HPLC analysis of *R,R*-**2**:*S,S*-**2**:*S,S*-**1** (1:1:2 molar ratio, 5 mM total concentration) library in the presence of 1 M $NaNO_3$.

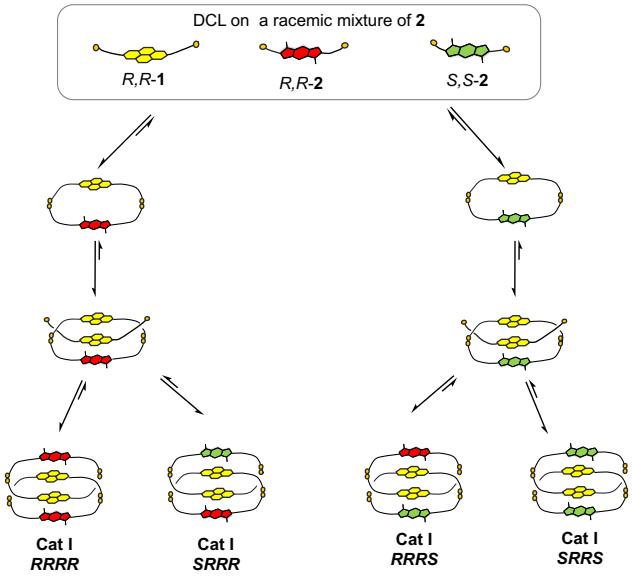


Figure 2.37. Schematic representation (simplified) of the pathways leading to **Cat I** isomers.

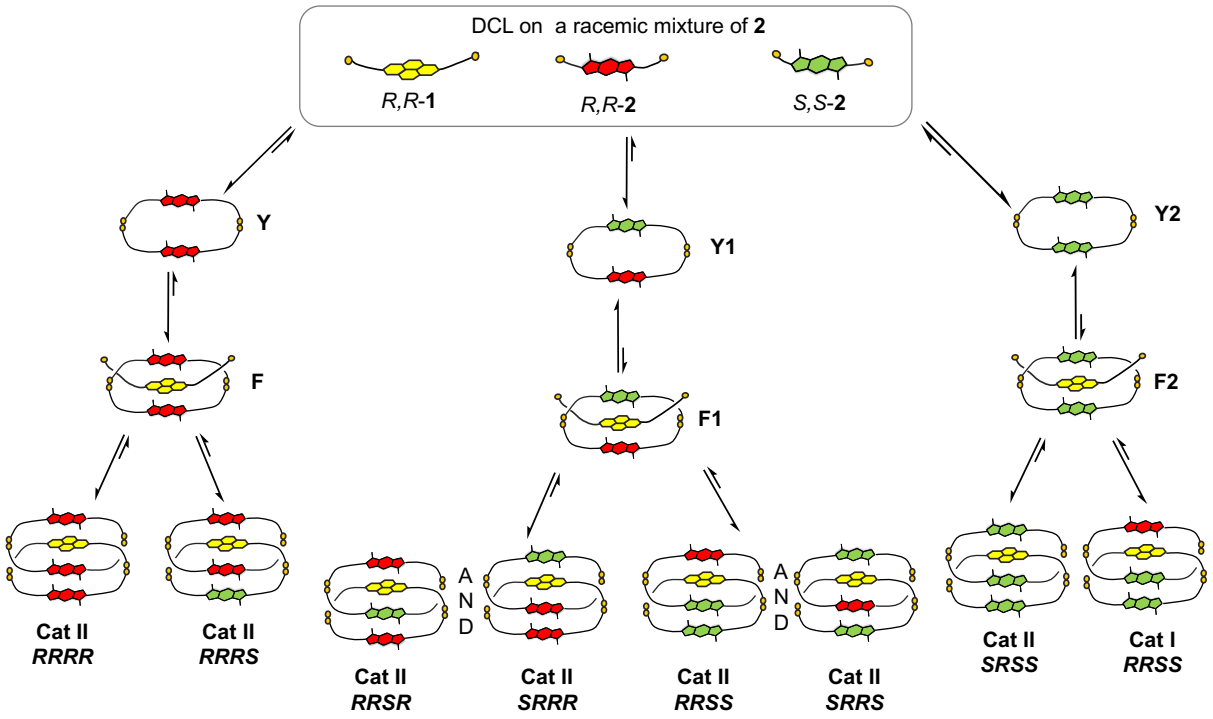


Figure 2.38. Schematic representation (simplified) of the pathways leading to **Cat II** isomers.

The mechanism through which structural divergence is achieved in this case follows a different path compared to the one for the NDI racemic mixture library. The first divergence point is still present (through the formation of **Cat I *RRRR***), so the homochiral heterodimer **X** is thermodynamically favoured in this DCLs as was in the previous case. The weak stability of the heterochiral heterodimer (formed from *R,R*-**1** and *S,S*-**2**) remains the same in comparison to **X**. The stabilities of the cyclic homodimers (**Y**, **Y1** and **Y2**), chirality mismatch and supramolecular interaction play a role in the SDRRM outcome. The PM7 energy calculations (Table 2.2) of all the catenanes in Figure 2.38 are in accordance with the experimental data: four catenanes formed in good yields and three other catenanes in low yields. In the SDRRM based on a racemic mixture of the NDI building block, *R,R*-**2** (BDT) discriminates between the two heterodimers, and **Y** (*R,R*-**2** dimer). In the present case, SDRRM based on a racemic mixture of BDT building block, *R,R*-**1** (NDI) does not discriminate between the three BDT dimers, leading to a multitude of catenanes. Another difference between the first and second SDRRMs is that, in the latter case, other molecules can be

threaded through the BDT dimer, while the corresponding NDI dimer present in the first SDRRM is too tight to allow any threading.

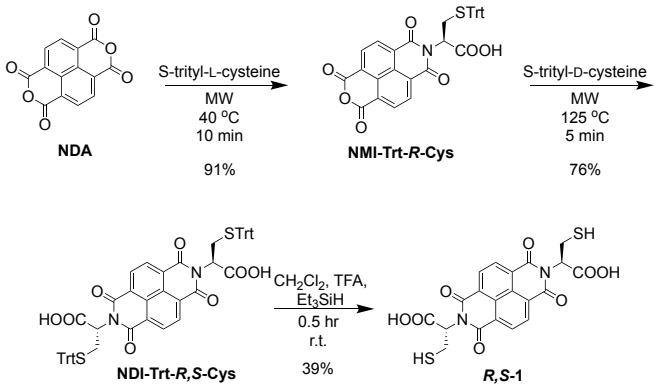
Table 2.2. Energies of different types of **Cat II**.

Catenane	Energy (kJ mol^{-1})
DADD **Cat II *RRRR***	-8501.01
DADD **Cat II *SRRR***	-8464.48
DADD **Cat II *RRSR***	-8530.06
DADD **Cat II *RRRS***	-8521.01
DADD **Cat II *SRSR***	-8492.47
DADD **Cat II *SRRS***	-8488.25
DADD **Cat II *RRSS***	-8555.99
DADD **Cat II *SRSS***	-8522.36

To sum up this discussion, the BDT racemic mixture DCL gives a less pronounced structural divergent reaction. For **Cat I**, only two isomers are formed, while for **Cat II**, seven out of eight possible isomers were selected. Based on this, this system is not as selective as the previous one. *R,R*-**1** does not discriminate between the BDT dimers, hence the formation of multiple **Cat II** isomers. The system is more selective in the case of **Cat I** as one of the species is formed in 42% yield. The tightness of the dimers (NDI *vs*. BDT) also plays a role as the NDI dimer does not allow any threading.

2.7. Analysis of the library containing *R,S*-1, and *R,R*-2

The last part of the discussion about NDI – BDT system covers the behaviour of a DCL containing the *meso*-NDI (*R,S*-**1**) and *R,R*-**2**. This NDI has on one side an L-cysteine group and, on the other side, a D-cysteine scaffold. The synthetic route started with the synthesis of a naphthalenemonoimide (NMI) from NDA and *S*-trityl-L-cysteine in 1:1 ratio. The NMI was further reacted with *S*-trityl-D-cysteine to obtain a desymmetrised NDI with cysteine of opposite chirality on each side. The trityl deprotection afforded *R,S*-**1** (Scheme 2.4).



Scheme 2.4. The synthetic scheme for *R*,*S*-**1**.

The DCL was set of by dissolving the building blocks in 1:1 ratio in the presence of 1 M $NaNO_3$. The analysis of the library indicated the formation of four [2]catenanes, which was unexpected: three types of **Cat I** catenanes and one type of **Cat II** catenane. Unfortunately, the yield of each catenane could not be determined by integrating the chromatogram because of the poor separation even after arduous method development. (Figures 2.39 and 2.40 – Figure 2.40 is an expanded version of 2.39 in which the yield of the catenanes is depicted in function of the other catenanes; the total relative yield of all catenanes is 100%). Despite this, the question that arises now is why are there four [2]catenanes formed? There are three possible isomers for **Cat I**: *RRRR-RSSR*, *RSRR-RRSR* and *RSSR-RRRR*, identified by the cysteine chirality on the left and right hand side of the catenane, as depicted in Figure 2.41 (for a Chemdraw® representation of *RRRR-RRRR,* see Figure 2.7). The heterodimer has directionality, and hence, the first and last isomers are different (Figure 2.41, top row). The steric and electrostatic repulsions are playing a role in the assembly, and all three isomers are formed in different proportions.

A different situation is observed for **Cat II** isomers. Only two isomers are theoretically possible: *RRRR-RSRR* and *RSRR-RRRR*. However, because **Y** (or *R,R*-**2** dimer) has no directionality, these two isomers are identical. This is in line with the HPLC analysis of the corresponding library. Experiments with the competitors (NDI-*S,S*-serine and NDI-*R,R*-serine) have led to a decrease in yields of the catenanes, as shown in Figure 2.39. It is interesting that the same behaviour as in the case of NDI racemic mixture DCL is replicated by the DCL containing *meso*-NDI (*R*,*S*-**1**) and *R,R*-**2**.

The stereocentre inversion has a significant influence on the library distribution, especially in the number of catenanes formed. For the *DAAD* catenane with two chiral centres inverted, three conformational isomers are observed. For the *DADD* catenane with one chiral centre inverted, only one conformational isomer is possible. This reinforces the steric and electrostatic repulsions between the carboxylate anions when the chirality changes. This has been assigned as one of the driving forces behind the structural divergence reactions through dynamic combinatorial chemistry described herein.

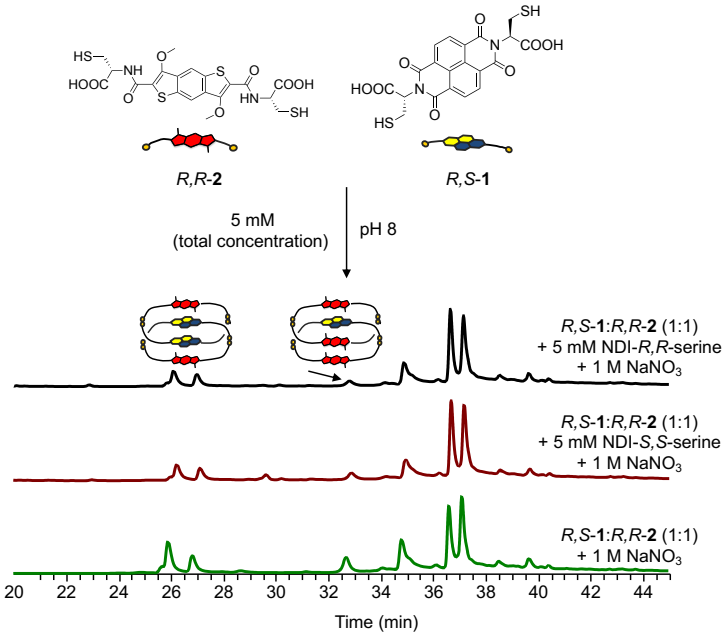


Figure 2.39. Reverse-phase HPLC analysis of *R,S*-**1**:*R,R*-**2** (1:1 molar ratio, 5 mM total concentration) library in the presence of 1 M $NaNO_3$, 1 M $NaNO_3$ + 5 mM NDI-*S,S*-serine and 1 M $NaNO_3$ + 5 mM NDI-*R,R*-serine as shown in the figure. Absorbance recorded at 389 nm. The unlabelled peaks did not ionise and could not be identified.

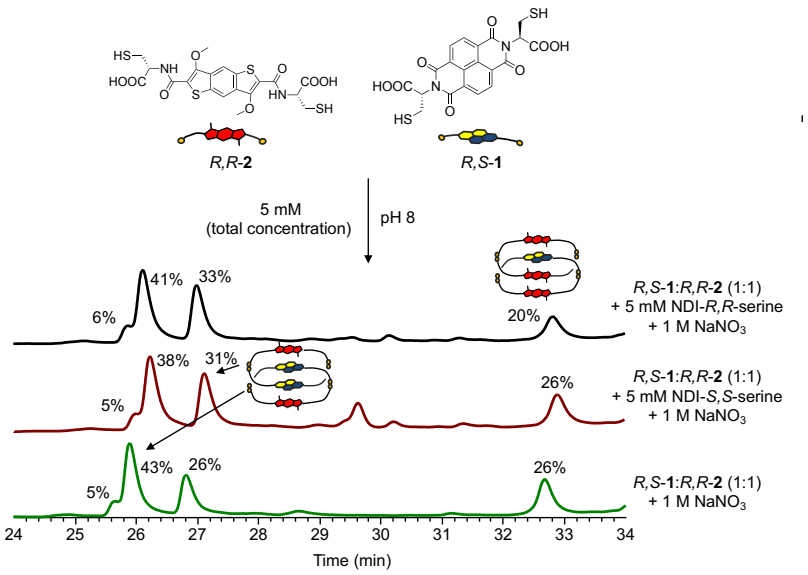


Figure 2.40. Reverse-phase HPLC analysis of *R,S*-**1**:*R,R*-**2** (1:1 molar ratio, 5 mM total concentration) library in the presence of 1 M $NaNO_3$, 1 M $NaNO_3$ + 5 mM NDI-*S,S*-serine and 1 M $NaNO_3$ + 5 mM NDI-*R,R*-serine as shown in the figure. Absorbance recorded at 389 nm. The unlabelled peaks did not ionise and could not be identified. Expanded version of Figure 2.39 with the yields for each catenane.

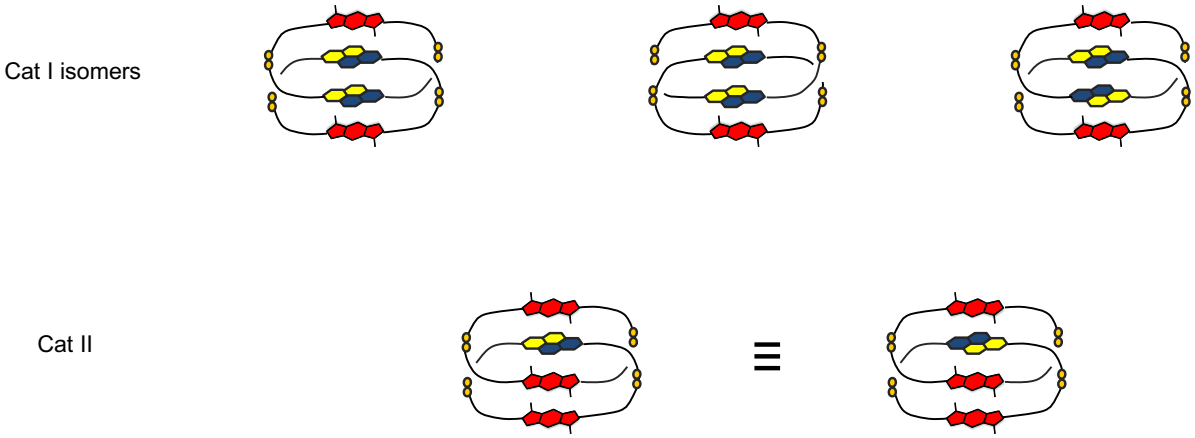


Figure 2.41. All possibilities of **Cat I** and **Cat II** isomers containing *R,S*-**1** (the yellow part represents the side with *R*-cysteine, and the blue part represents the side with *S*-cysteine).

The last part of this discussion briefly describes another structural divergence reaction in DCC using different building blocks. This work was performed by an MChem student, Kornelijus Balnius, under my direct supervision and guidance.

For this work, we have chosen rigid and flexible building blocks and DCLs were set up having the previously introduced NDI (*R,R*-**1** and *S,S*-**1**) as the π-acceptor component. Both building blocks (*i.e.* the rigid and flexible ones) have been previously studied in Sanders' and Pantoș' groups, but only in homochiral DCLs. The rigid molecule is a 1,5-dialkoxynaphthalene derivative (*R,R* or *S,S*-**1,5-DN**), while the flexible one is a 2,6-dialkoxynaphthalene derivative (*R,R* or *S,S*-**2,6-DN**) – Figure 2.42. These DN derivatives (both 1,5 and 2,6) are π-donor molecules. A full description of their behaviour as building blocks in homochiral DCLs can be found either in published papers or in a PhD thesis.[9,15–17]

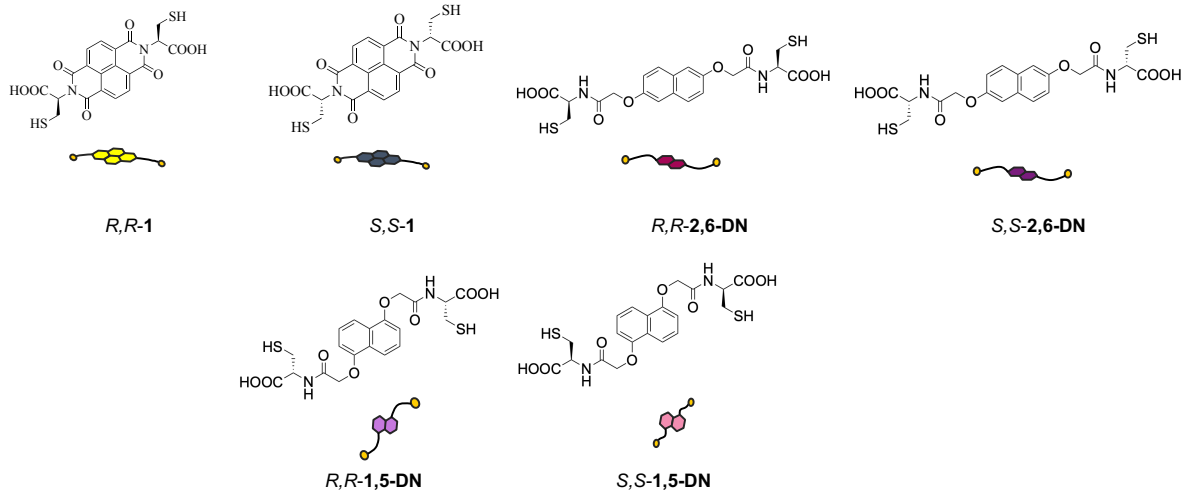


Figure 2.42. The structures and the cartoon representations of the molecules used to generate heterochiral DN-NDI-based DCLs.

2.8. Analysis of the libraries containing *R,R*-1/*S,S*-1 and *R,R*-1,5-DN/*S,S*-1,5-DN (homo, heterochiral and racemic libraries)

The distribution of the library containing the rigid DN is rather simple for both homochiral (*R,R*-**1** with *R,R*-**1,5-DN** or *S,S*-**1** with *S,S*-**1,5-DN**) and heterochiral (*R,R*-**1** with *S,S*-**1,5-DN** and *S,S*-**1** with *R,R*-**1,5-DN**) DCLs. The only molecule formed is a cyclic dimer (either homo or heterochiral) regardless the presence or absence of salt. When the racemic mixture library was analysed, the results showed the presence of only two major species: hetero and homochiral heterodimers (Figures 2.43 – 2.46). There is no other interaction between the components due to the extensive rigidity of the system. This can be described as a self-sorting process.

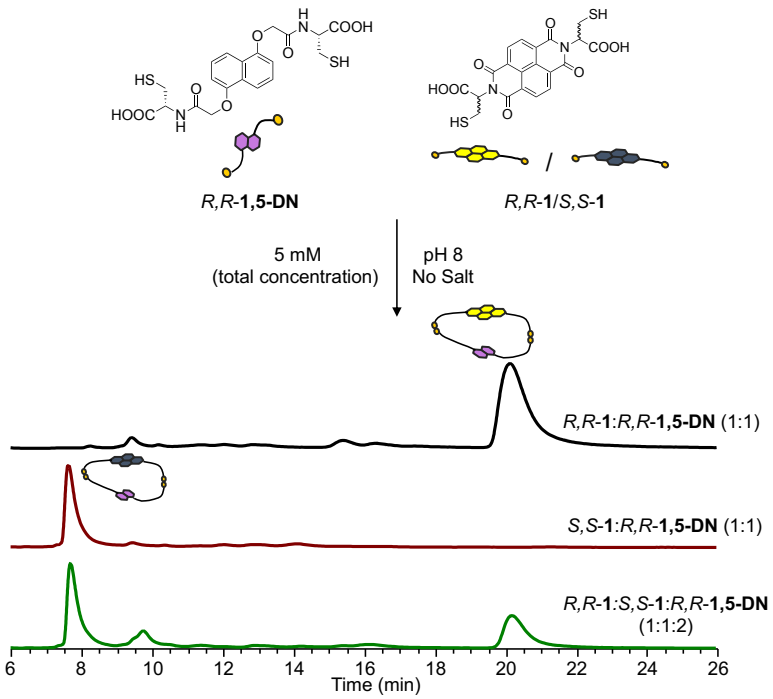


Figure 2.43. Reverse-phase HPLC analysis of *R,R*-**1**:*R,R*-**1,5-DN**, (1:1 molar ratio, 5 mM total concentration), *S,S*-**1**:*R,R*-**1,5-DN** (1:1 molar ratio, 5 mM total concentration), *R,R*-**1**:*S,S*-**1**:*R,R*-**1,5-DN** (1:1:2 molar ratio, 5 mM total concentration) libraries without salt.

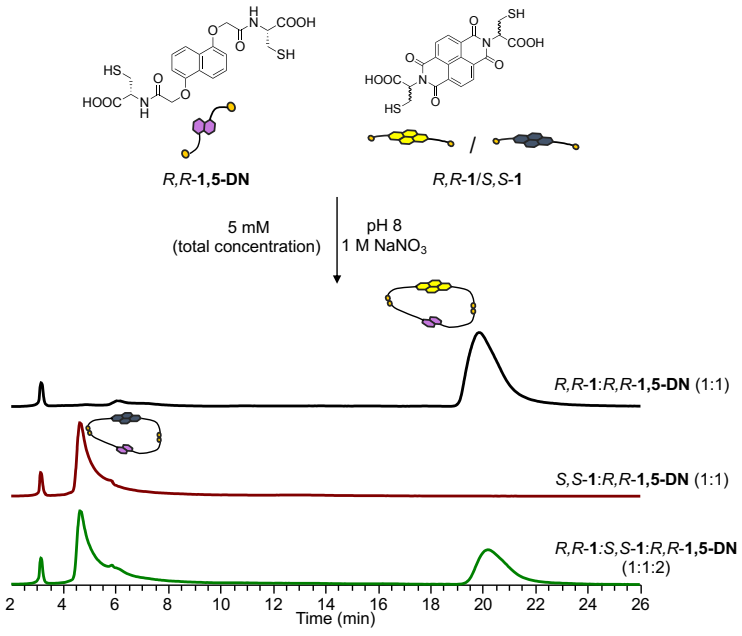


Figure 2.44. Reverse-phase HPLC analysis of *R,R*-**1**:*R,R*-**1,5-DN**, (1:1 molar ratio, 5 mM total concentration), *S,S*-**1**:*R,R*-**1,5-DN** (1:1 molar ratio, 5 mM total concentration), *R,R*-**1**:*S,S*-**1**:*R,R*-**1,5-DN** (1:1:2 molar ratio, 5 mM total concentration) libraries in the presence of 1 M $NaNO_3$.

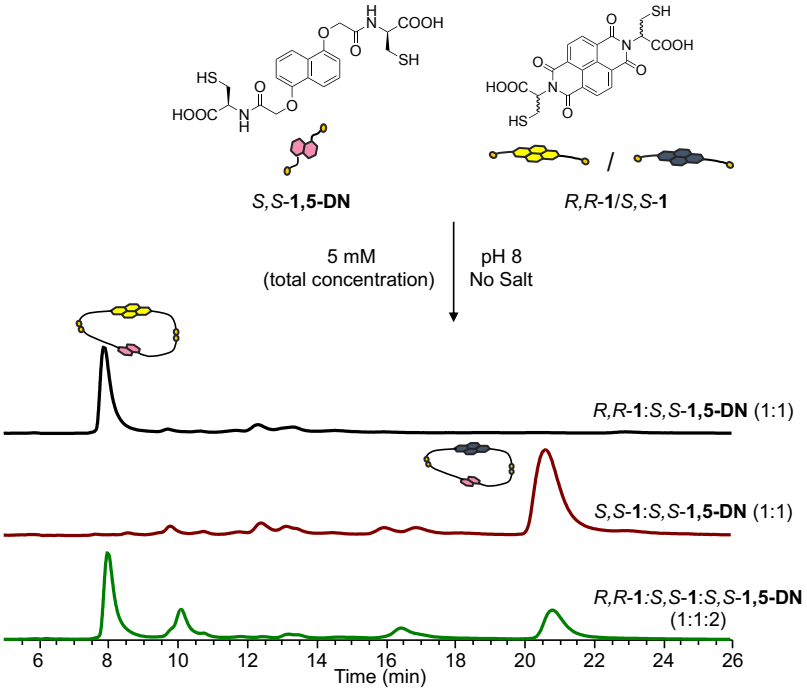


Figure 2.45. Reverse-phase HPLC analysis of *R,R*-**1**:*S,S*-**1,5-DN**, (1:1 molar ratio, 5 mM total concentration), *S,S*-**1**:*S,S*-**1,5-DN** (1:1 molar ratio, 5 mM total concentration), *R,R*-**1**:*S,S*-**1**:*S,S*-**1,5-DN** (1:1:2 molar ratio, 5 mM total concentration) libraries without salt.

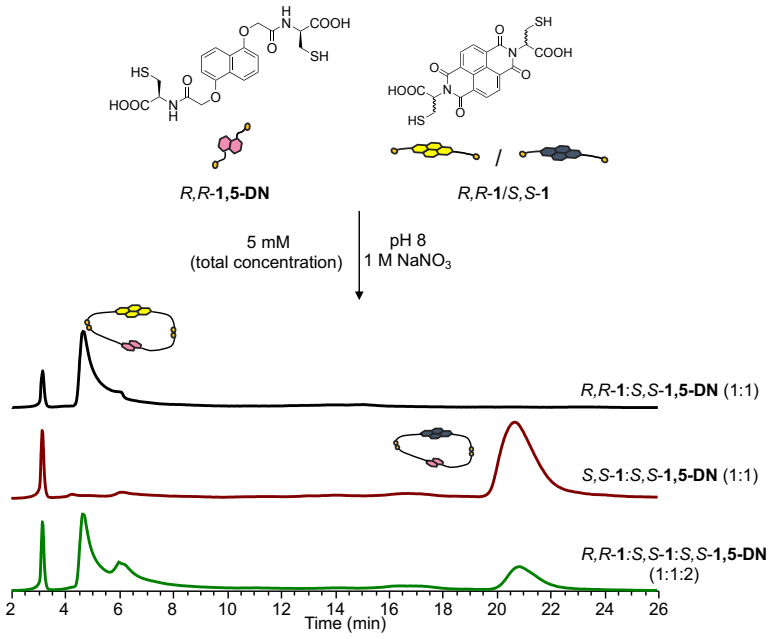


Figure 2.46. Reverse-phase HPLC analysis of *R,R*-**1**:*S,S*-**1,5-DN**, (1:1 molar ratio, 5 mM total concentration), *S,S*-**1**:*S,S*-**1,5-DN** (1:1 molar ratio, 5 mM total concentration), *R,R*-**1**:*S,S*-**1**:*S,S*-**1,5-DN** (1:1:2 molar ratio, 5 mM total concentration) libraries in the presence of 1 M $NaNO_3$.

2.9. Analysis of the libraries containing *R,R*-1/*S,S*-1 and *R,R*-2,6-DN/*S,S*-2,6-DN (homo-, heterochiral and racemic libraries)

The second system contains the flexible DN and NDI as building blocks. The homochiral library (*R,R*-**1** with *R,R*-**2,6-DN** or *S,S*-**1** with *S,S*-**2,6-DN**) are similar to the one previously published.[13] The libraries without salt produce only dimers as major components. When 1 M $NaNO_3$ is added, the main products are two different [2]catenanes (with *DAAD* and *DADD* stacking) along with a heterodimer. In the case of the homochiral libraries, both types of catenanes are formed compared to the BDT-based DCLs. This indicates that the 2,6-DN containing species are more flexible than the BDT counterparts and less selective towards catenanes' stacking differentiation. However, the distribution for the heterochiral libraries (*R,R*-**1** with *S,S*-**2,6-DN** and *S,S*-**1** with *R,R*-**2,6-DN**) is different. No catenane is formed, but only a heterochiral heterodimer. This highlights again that changing the chirality imposes steric and electrostatic repulsions, which lead to structural changes. In this particular case, the changing in chirality does not even allow for the threading. The chromatograms corresponding to the libraries with racemic mixture (both NDI and DN are used as racemates) indicate that building blocks with opposite chirality interact less than in the case of the BDT. Apart from common species forming in the hetero and homochiral DCLs, only a trimer (containing units with both chiralities) is observed (Figures 2.47 – 2.50). However, this is a "clean" structural divergent reaction, in which each enantiomer behaves almost independently of each other.

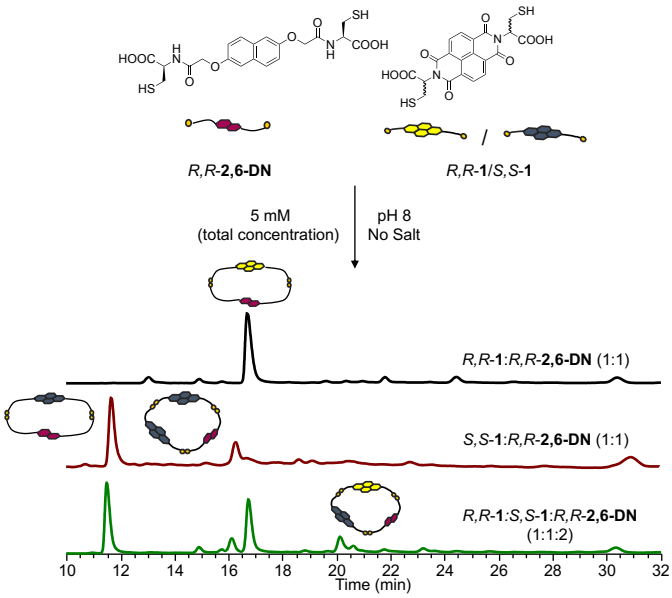


Figure 2.47. Reverse-phase HPLC analysis of *R,R*-**1**:*R,R*-**2,6-DN**, (1:1 molar ratio, 5 mM total concentration), *S,S*-**1**:*R,R*-**2,6-DN** (1:1 molar ratio, 5 mM total concentration), *R,R*-**1**:*R,R*-**1**:*S,S*-**2,6-DN** (1:1:2 molar ratio, 5 mM total concentration) libraries without salt.

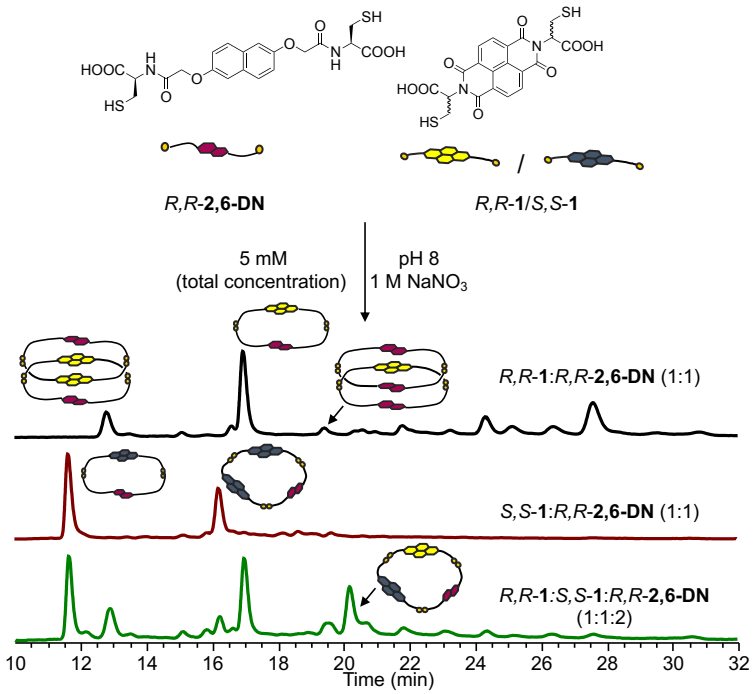


Figure 2.48. Reverse-phase HPLC analysis of *R,R*-**1**:*R,R*-**2,6-DN**, (1:1 molar ratio, 5 mM total concentration), *S,S*-**1**:*R,R*-**2,6-DN** (1:1 molar ratio, 5 mM total concentration), *R,R*-**1**:*S,S*-**1**:*R,R*-**2,6-DN** (1:1:2 molar ratio, 5 mM total concentration) libraries in the presence of 1 M $NaNO_3$.

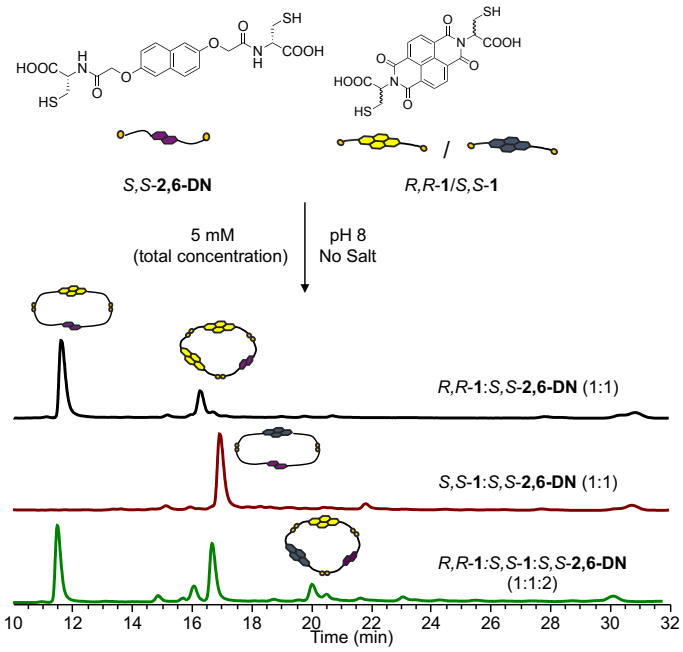


Figure 2.49. Reverse-phase HPLC analysis of *R,R*-**1**:*S,S*-**2,6-DN**, (1:1 molar ratio, 5 mM total concentration), *S,S*-**1**:*S,S*-**2,6-DN** (1:1 molar ratio, 5 mM total concentration), *R,R*-**1**:*S,S*-**1**:*S,S*-**2,6-DN** (1:1:2 molar ratio, 5 mM total concentration) libraries without salt.

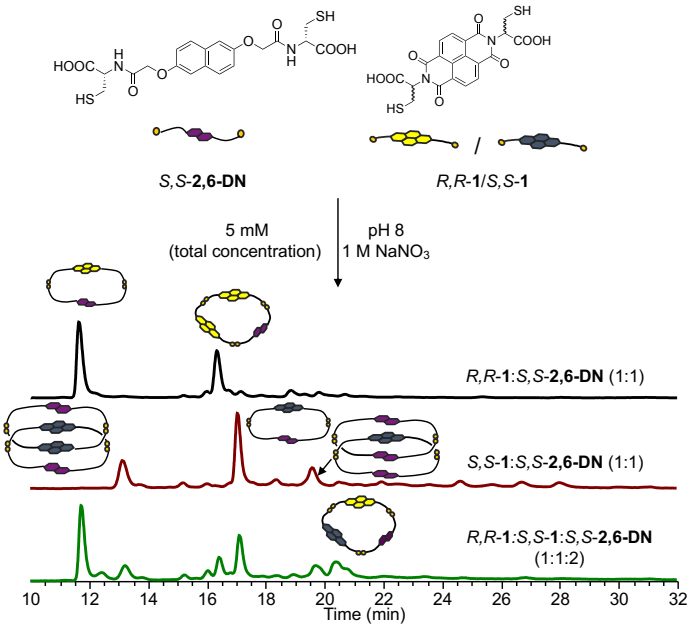


Figure 2.50. Reverse-phase HPLC analysis of *R,R*-**1**:*S,S*-**2,6-DN**, (1:1 molar ratio, 5 mM total concentration), *S,S*-**1**:*S,S*-**2,6-DN** (1:1 molar ratio, 5 mM total concentration), *R,R*-**1**:*S,S*-**1**:*S,S*-**2,6-DN** (1:1:2 molar ratio, 5 mM total concentration) libraries in the presence of 1 M $NaNO_3$.

2.10. Conclusion

This chapter has introduced and discussed the first ever structural divergent reactions on racemic mixtures using DCC. The first DCL system (*R,R*-**2** with racemic NDI) is the result of the thermodynamically favoured dimers, followed by stereospecific amplification of a certain pseudorotaxane. This highlights the importance of supramolecular recognition and chirality mismatch in controlling the formation of catenanes. In contrast, the next SDRRM (*R,R*-**1** with racemic BDT) gives another piece of information towards understanding supramolecular interactions. The chiral recognition is not so pronounced: for **Cat I**, two of three isomers are formed, while for **Cat II** out of eight possible isomers, only five are observed. This shows that *R,R*-**1** does not discriminate between the isomers of the BDT dimer. A *meso*-NDI (*R,S*-**1**) was also used, which introduced directionality in the *R,R*-**2**:*R,S*-**1** heterodimer, leading to the formation of stereoisomers of **Cat I**.

The use of rigid DN (the 1,5 isomer) has led to a system in which structural divergence is not present, but rather a self-sorting process is observed. This outcome must be due to the overall tightness of the heterodimers, which does not allow for larger structures to be formed. In contrast, the use of the flexible DN (the 2,6-isomer) allowed for the third SDRRM to be discovered. The chirality mismatch plays an important role, leading to only homochiral catenanes, whereas the heterochiral library yields just a dimer and a trimer.

DCLs take place in water and does not involve expensive chiral metal catalysts. All the work described in this chapter suggests the importance of stereospecific supramolecular recognition in Nature. The study shows that not all the DCC systems are prone to structural divergence, and the importance of the size of the cavity of the dimers also plays an essential role.

This work outlines some important aspects that lead to an SDRRM DCC behaviour:

- It is crucial that both building blocks are not rigid; there should be space in the one of the dimers cavity's to allow threading of another building block. This is necessary for avoiding the self-sorting process (*e.g.*, DCL of NDI and 1,5-DN case). This can be predictable using molecular modelling.

- The need for supramolecular interactions and formation of a stable pseudorotaxane: the NDI – 2,6-DN and NDI (racemic mixtures) – BDT cases are good examples.

2.11. Supporting figures and tables:

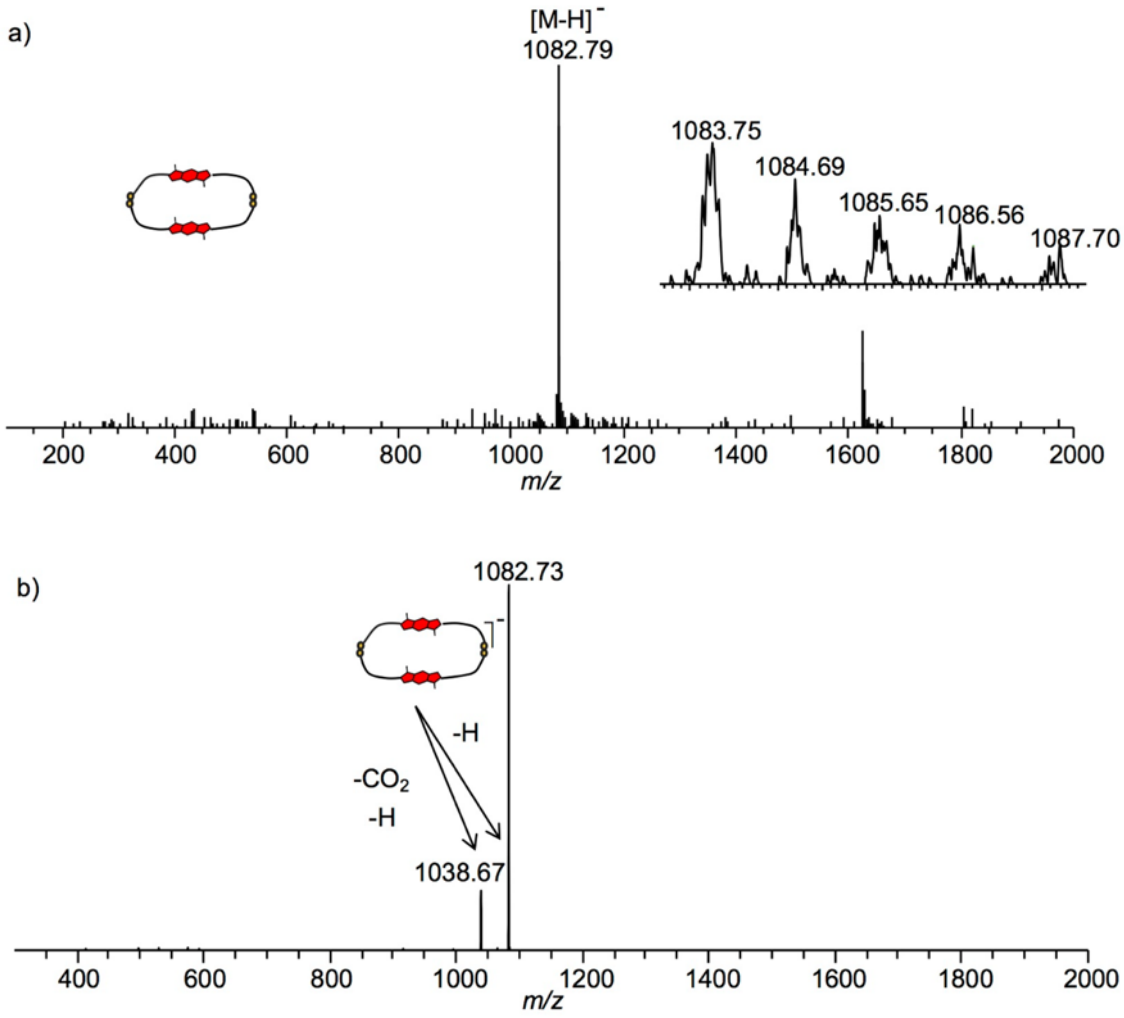


Figure 2.51. a) MS (-*ve*) of **Y**; zoom of molecular ion is shown as inset and b) MS/MS (-*ve*) of **Y**.

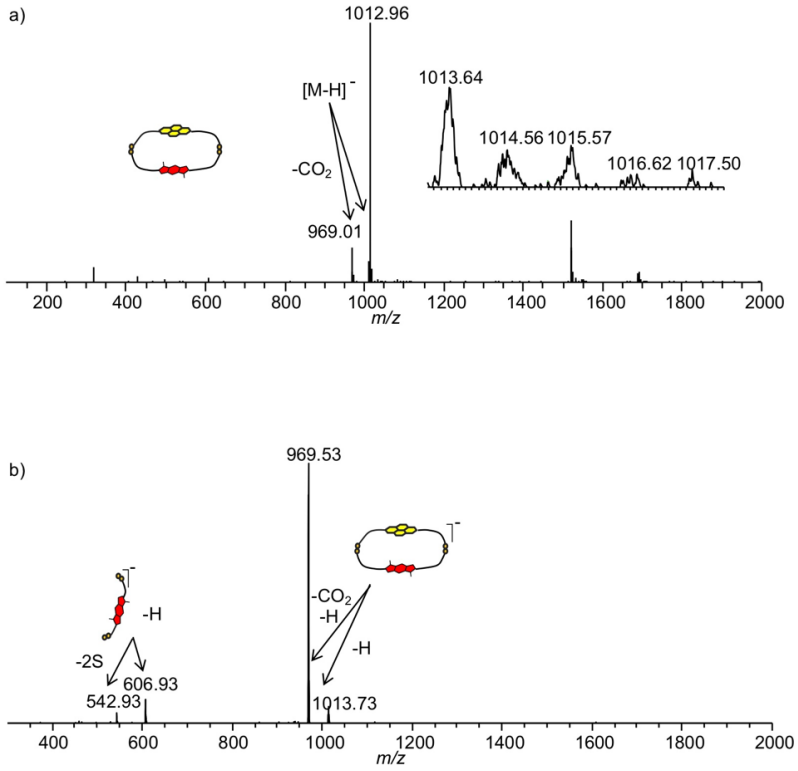


Figure 2.52. a) MS (-*ve*) of **X**; zoom of molecular ion is shown as inset and b) MS/MS (-*ve*) of **X**.

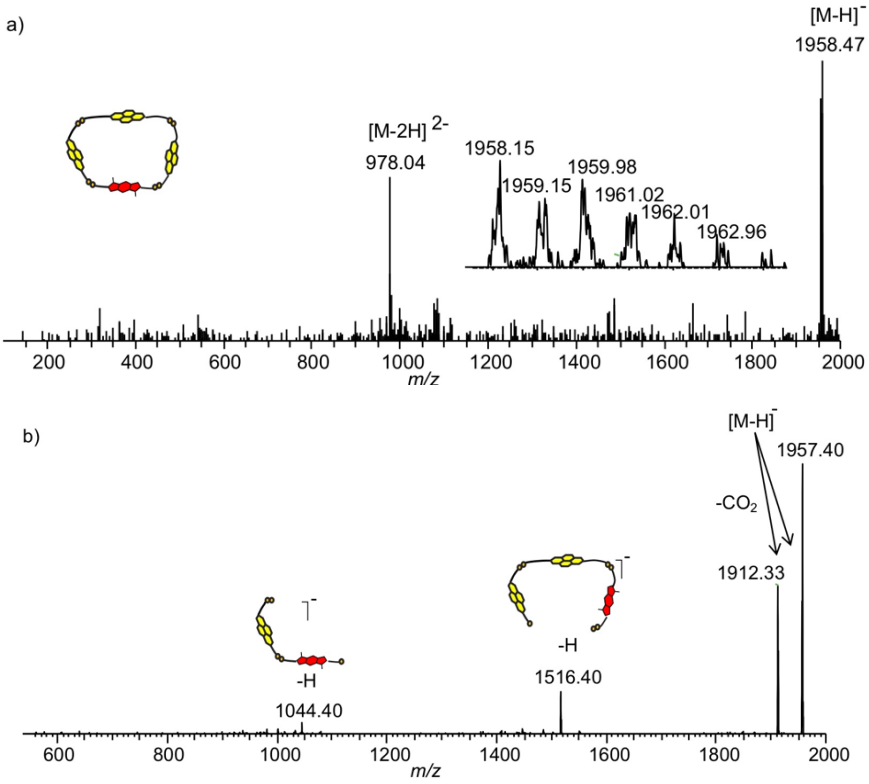


Figure 2.53. a) MS (-*ve*) of tetramer based on three *R,R*-**1** and one *R,R*-**2**; zoom of molecular ion is shown as inset and b) MS/MS (-*ve*) of tetramer.

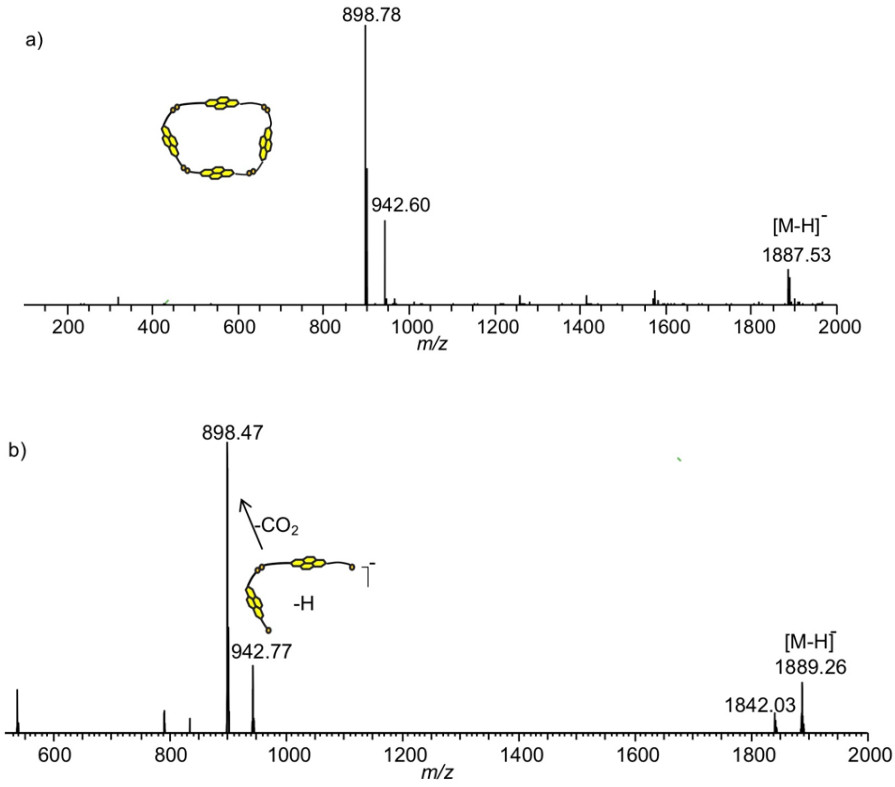


Figure 2.54. a) MS (-*ve*) of *R,R*-**1** tetramer; zoom of molecular ion is shown as inset and b) MS/MS (-*ve*) of tetramer.

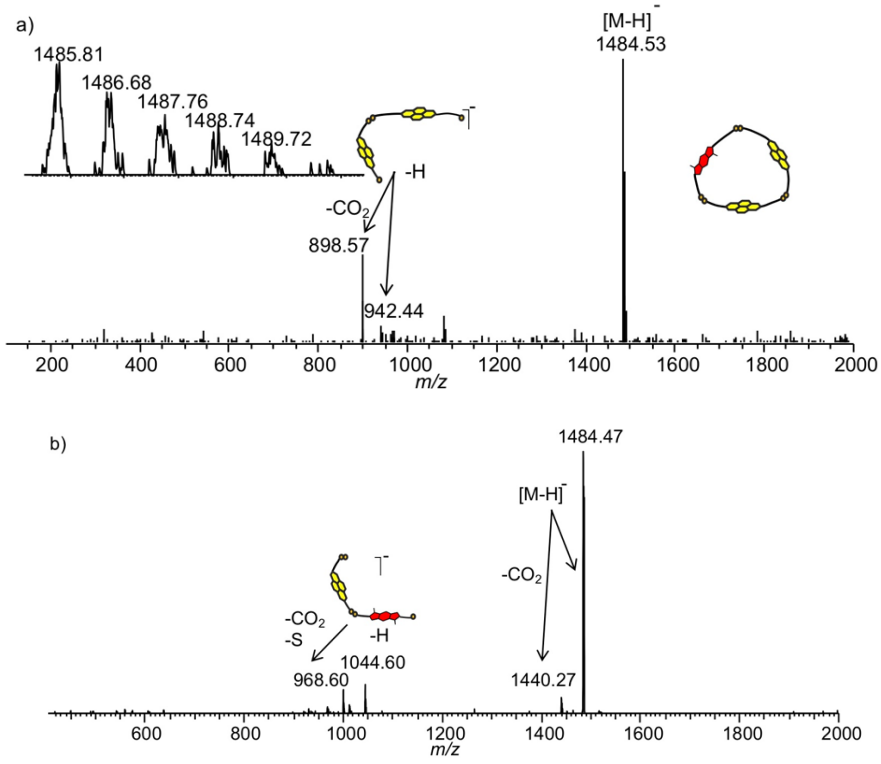


Figure 2.55. a) MS (-*ve*) of heterotrimer based on two *R,R*-**1** units and one *R,R*-**2** unit; zoom of molecular ion is shown as inset and b) MS/MS (-*ve*) of heterotrimer.

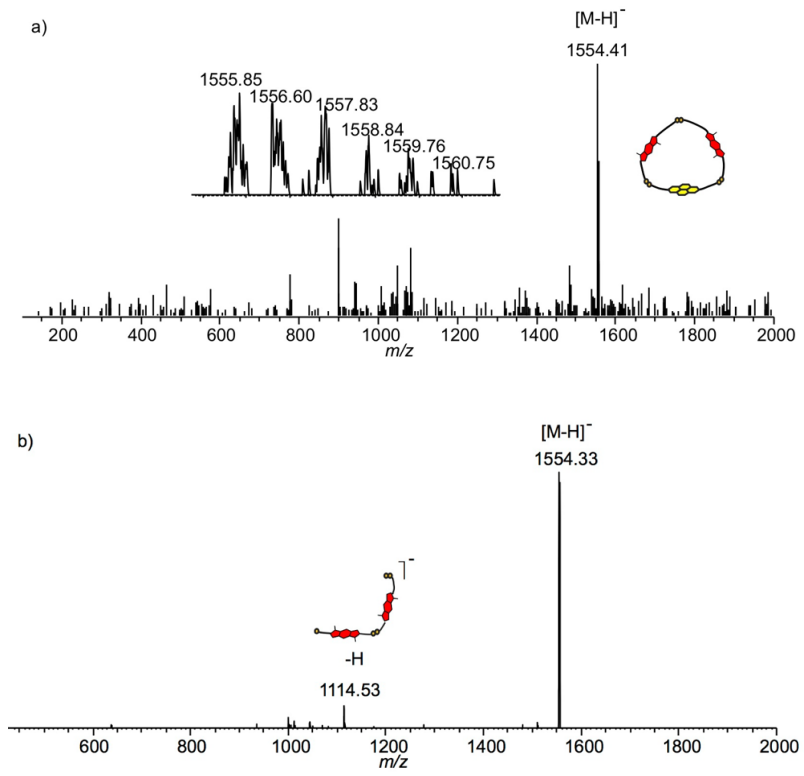


Figure 2.56. a) MS (-*ve*) of heterotrimer based on one *R,R*-**1** unit and two *R,R*-**2** unit; zoom of molecular ion is shown as inset and b) MS/MS (-*ve*) of heterotrimer.

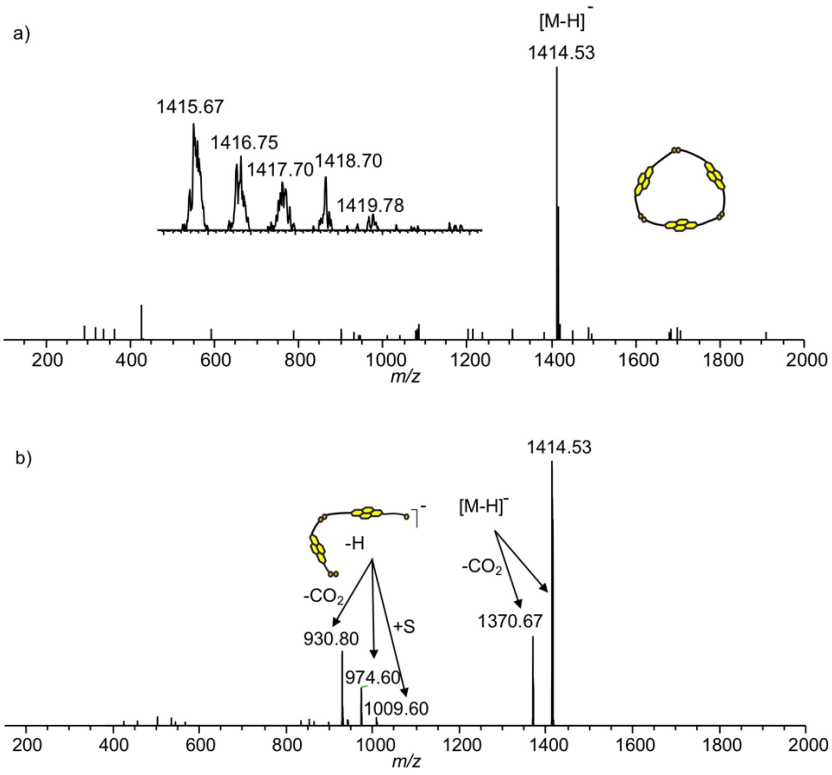


Figure 2.57. a) MS (-*ve*) of *R,R*-**1** homotrimer; zoom of molecular ion is shown as inset and b) MS/MS (-*ve*) of homotrimer.

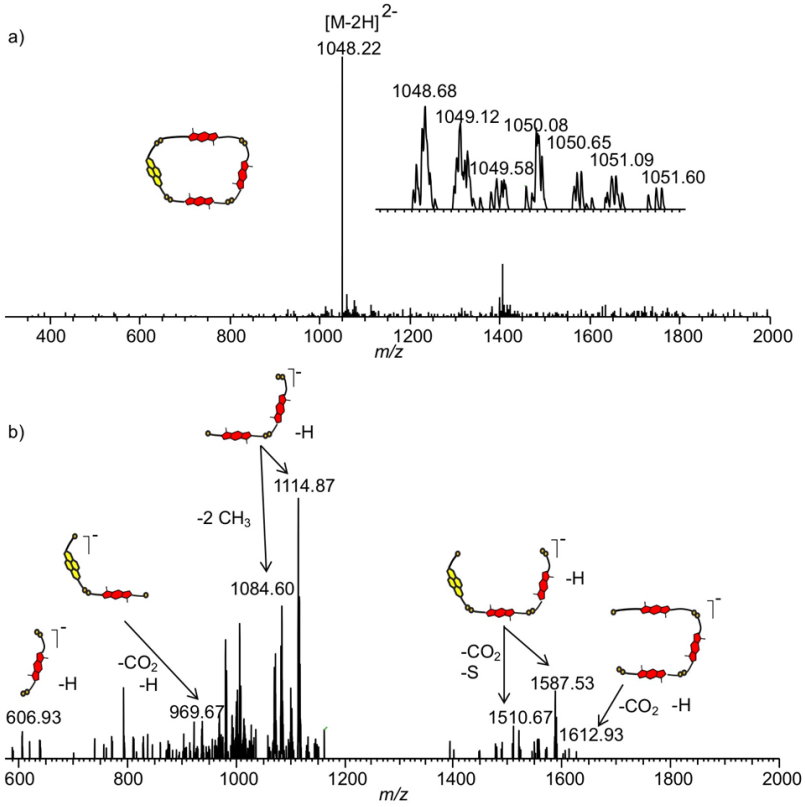


Figure 2.58. a) MS (-*ve*) of heterotetramer based on one *R,R*-**1** unit and three *R,R*-**2** unit; zoom of molecular ion is shown as inset and b) MS/MS (-*ve*) of heterotetramer.

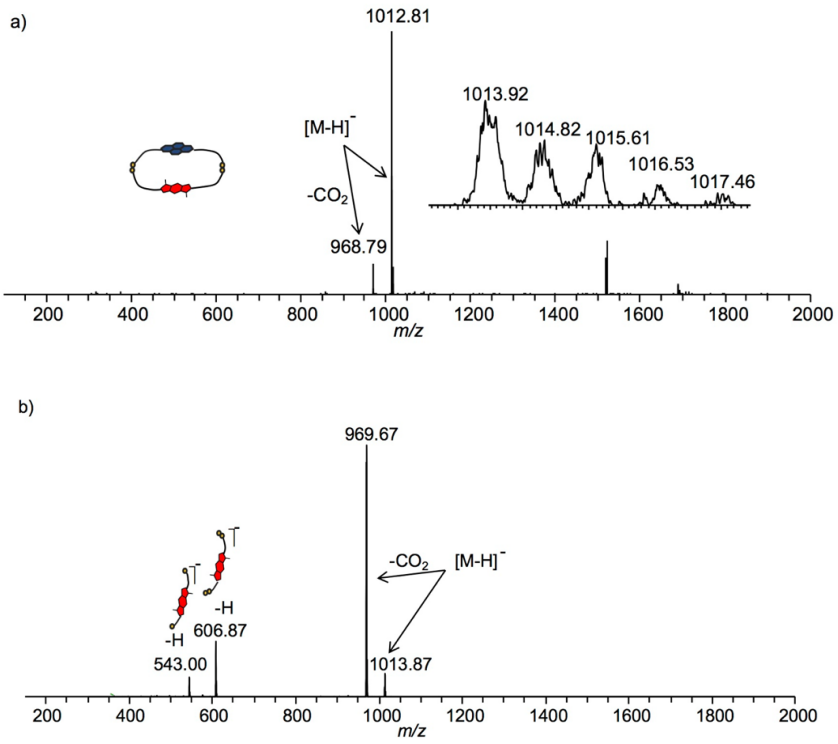


Figure 2.59. a) MS (-*ve*) of **Z**; zoom of molecular ion is shown as inset and b) MS/MS (-*ve*) of **Z**.

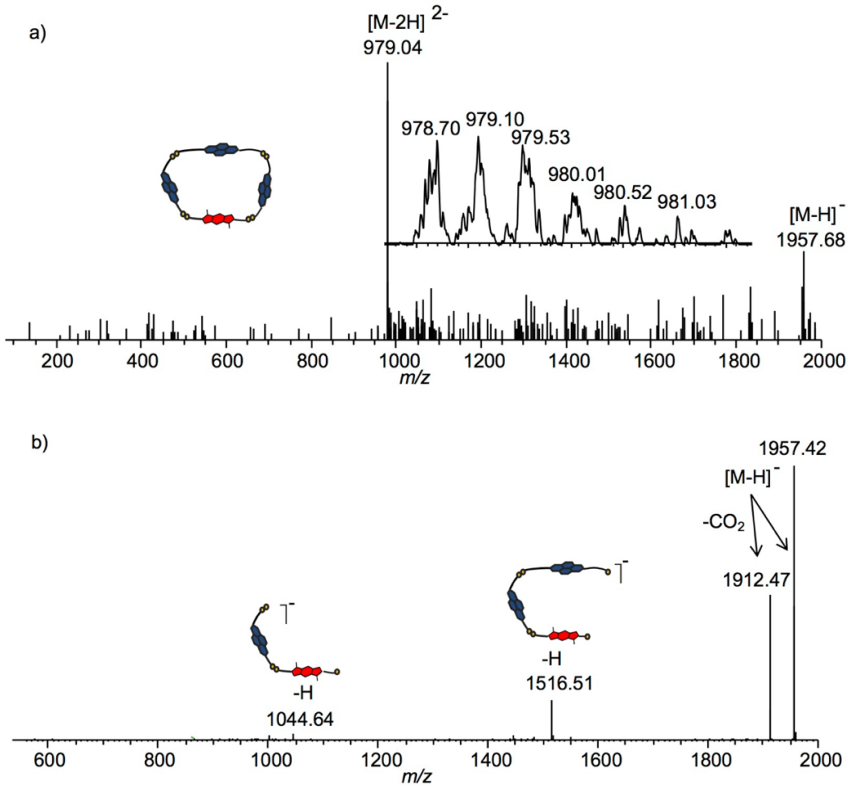


Figure 2.60. a) MS (-*ve*) of heterotetramer based on three *S,S*-**1** units and one *R,R*-**2** unit; zoom of molecular ion is shown as inset and b) MS/MS (-*ve*) of heterotetramer.

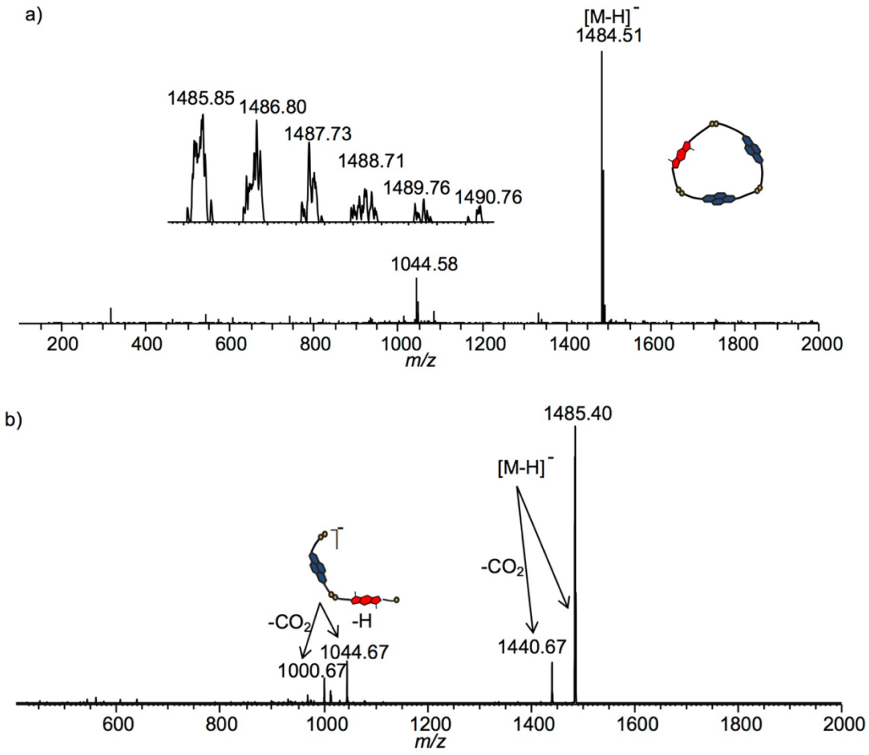


Figure 2.61. a) MS (-*ve*) of heterotrimer based on two *S,S*-**1** units and one *R,R*-**2** unit; zoom of molecular ion is shown as inset and b) MS/MS (-*ve*) of heterotrimer.

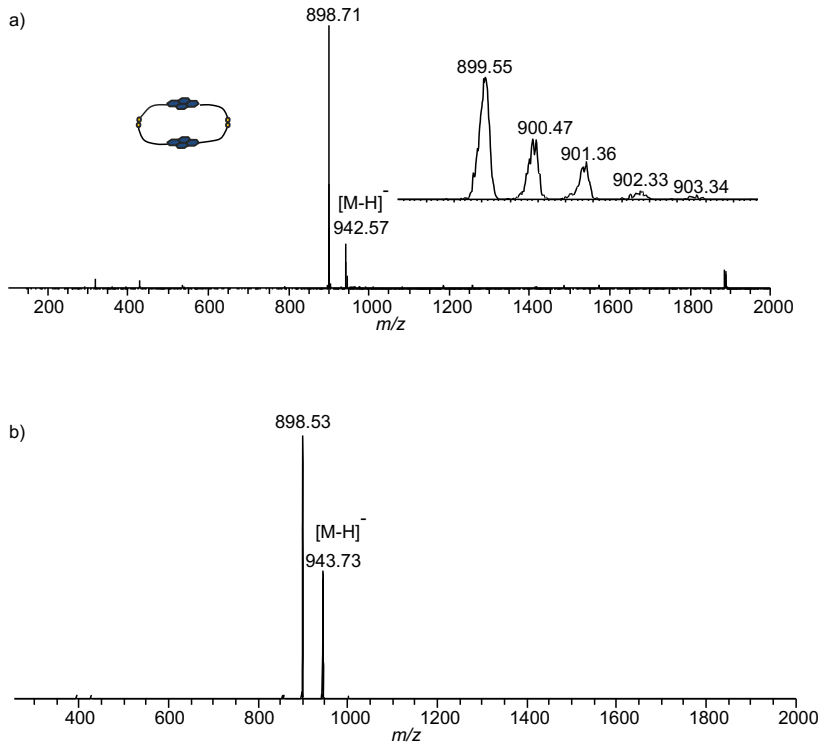


Figure 2.62. a) MS (-*ve*) of *S,S*-**1** homodimer; zoom of molecular ion is shown as inset and b) MS/MS (-*ve*) of homodimer.

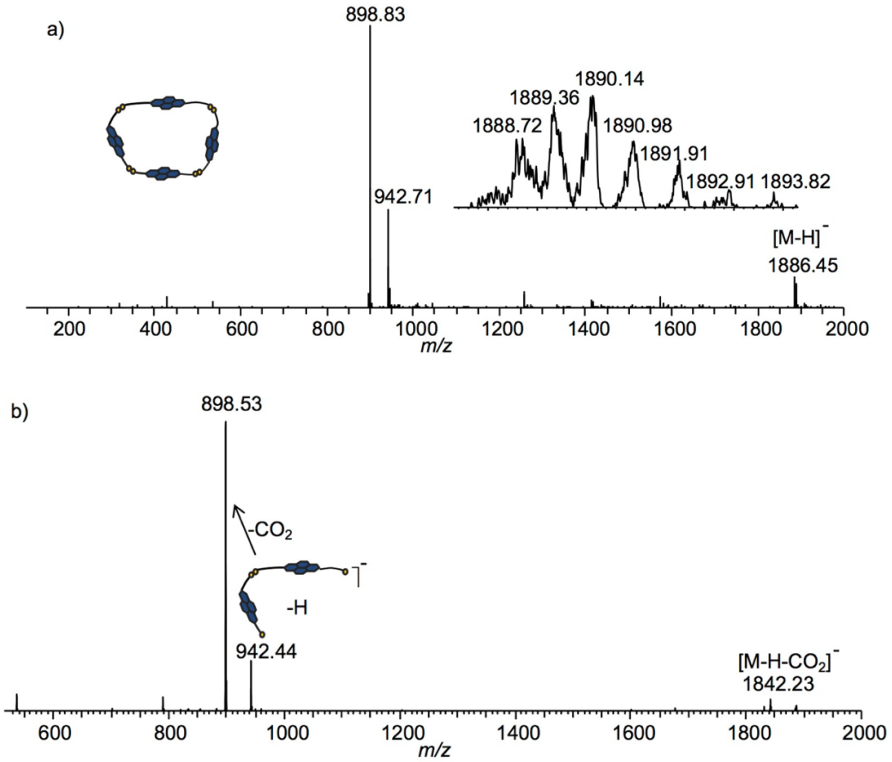


Figure 2.63. a) MS (-*ve*) of *S,S*-**1** homotetramer; zoom of molecular ion is shown as inset and b) MS/MS (-*ve*) of homotetramer.

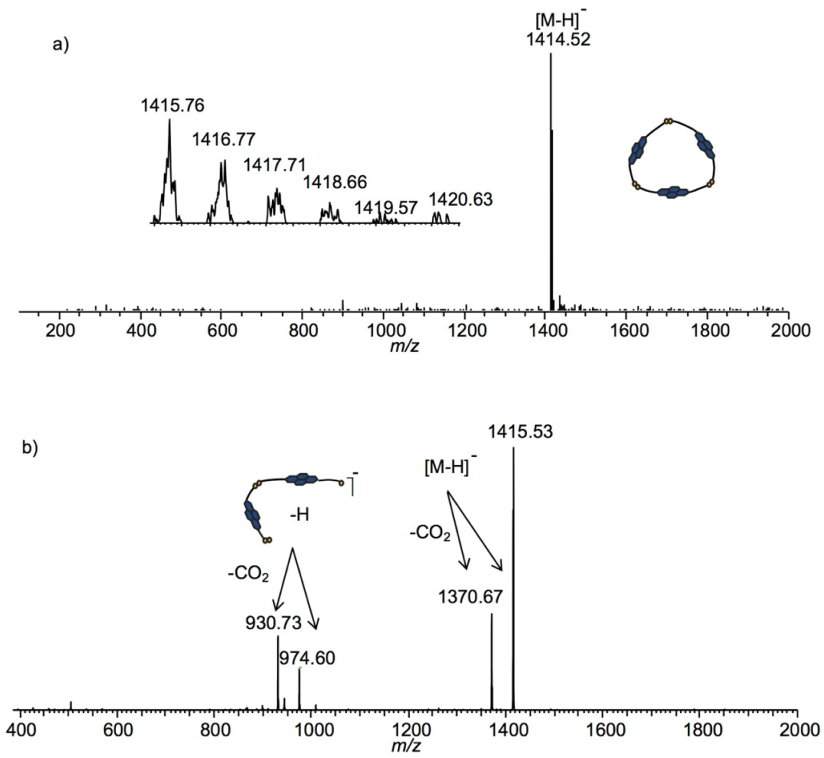


Figure 2.64. a) MS (-*ve*) of *S,S*-**1** homotrimer; zoom of molecular ion is shown as inset and b) MS/MS (-*ve*) of homotrimer.

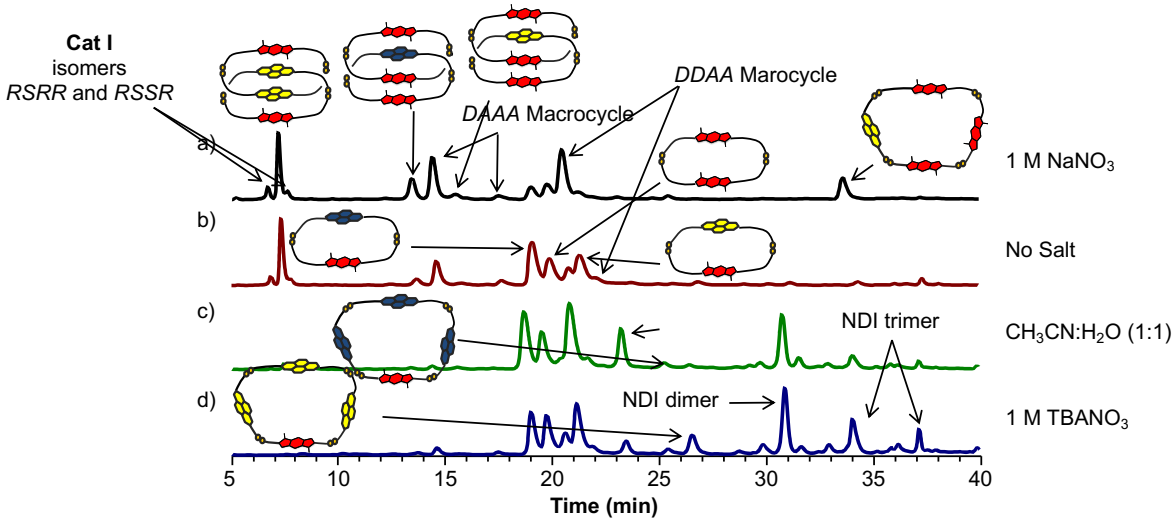


Figure 2.65. Reverse-phase HPLC analyses of *R,R*-**2**:*R,R*-**1**:*S,S*-**1** (2:1:1 molar ratio, 5 mM total concentration) library a) in the presence of 1 M $NaNO_3$, b) no salt, c) in a mixture of CH_3CN:H_2O (1:1) and d) in the presence of 1 M $TBANO_3$. Absorbances recorded at 389 nm. The unlabelled peaks did not ionise and could not be identified.

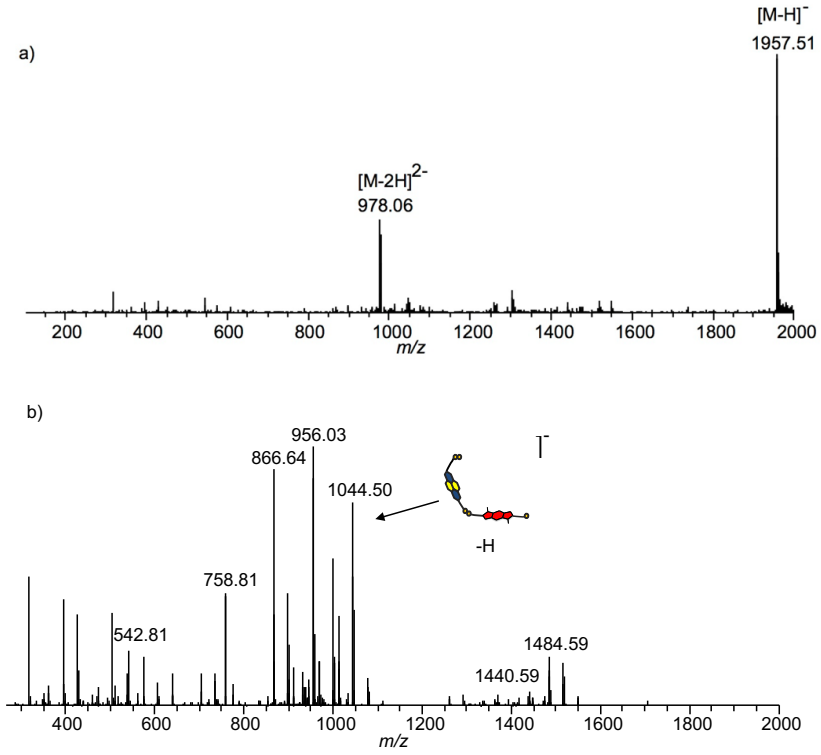


Figure 2.66. a) MS (-*ve*) of heterotetramer based on three *R,R*-**1**/*S,S*-**1** units and one *R,R*-**2** unit and b) MS/MS (-*ve*) of heterotetramer.

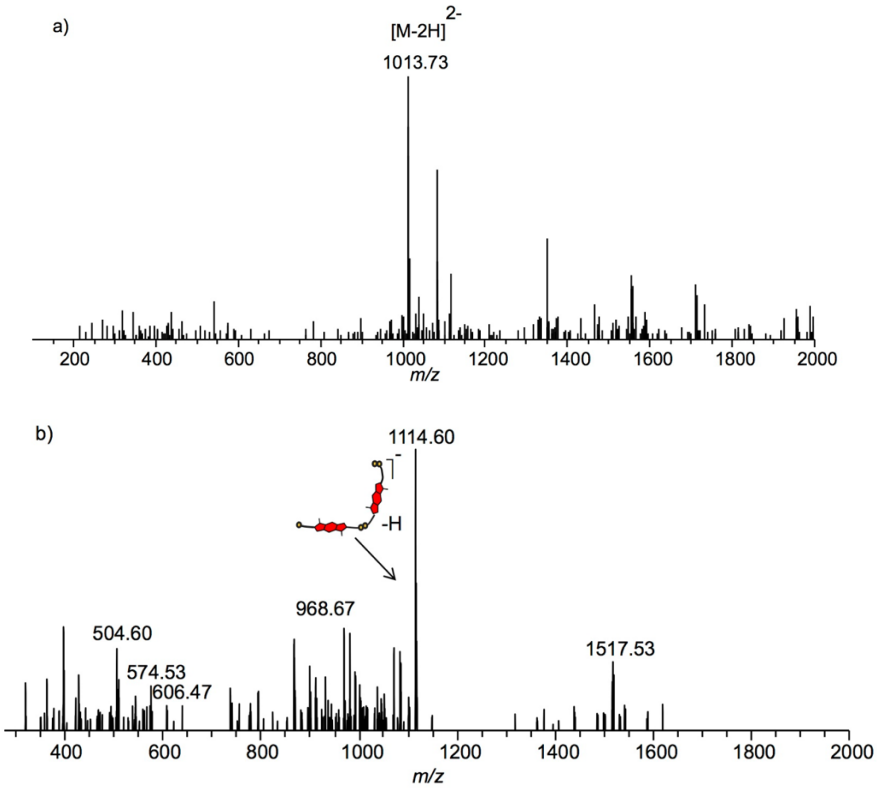


Figure 2.67. a) MS (-*ve*) of heterotetramer based on two *R,R*-**1**/*S,S*-**1** units and two *R,R*-**2** units and b) MS/MS (-*ve*) of heterotetramer.

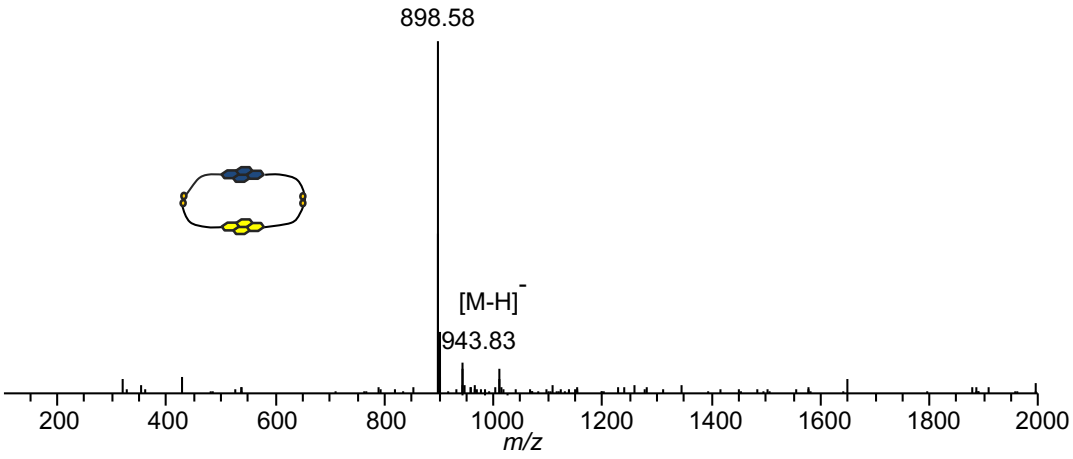


Figure 2.68. MS (-*ve*) of heterochiral NDI dimer.

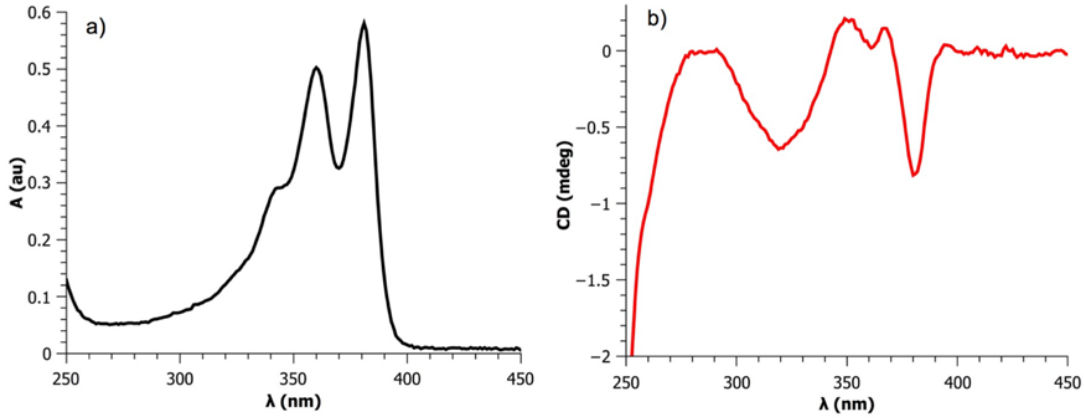


Figure 2.69. a) UV-Vis spectrum and b) CD spectrum of *R,R*-**1** at 23 °C in CH_3CN / H_2O mixture (66.66:33.34 *v/v*).

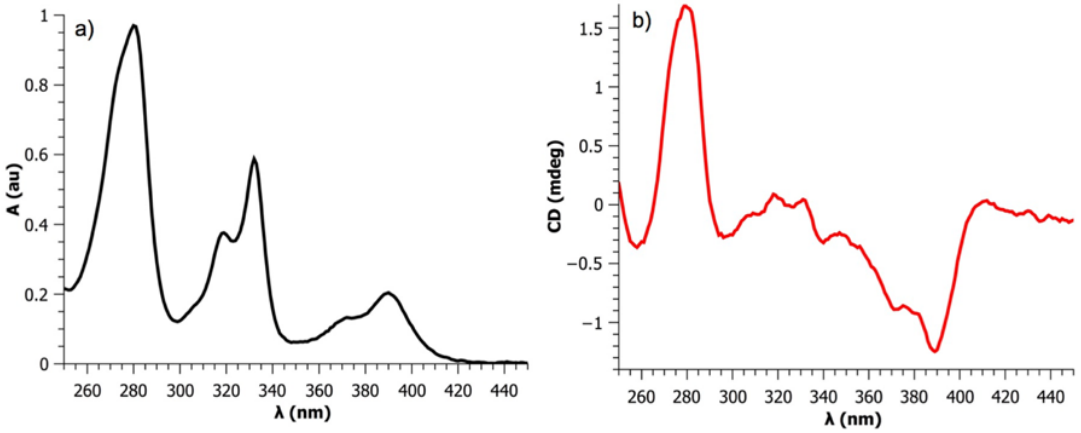


Figure 2.70. a) UV-Vis spectrum and b) CD spectrum of *R,R*-**2** at 23 °C. CH_3CN / H_2O mixture (66.66:33.34 *v/v*).

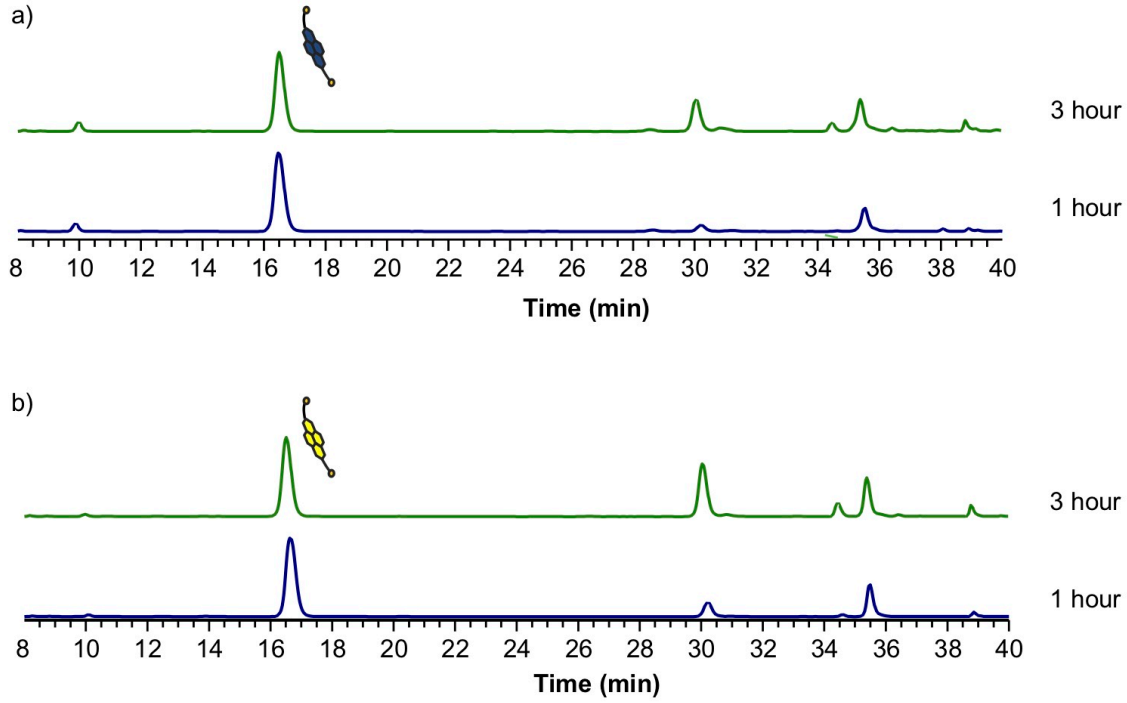


Figure 2.71. Reverse-phase HPLC analysis of a) *S,S*-**1** (2.5 mM total concentration) after 1 and 3 hours and b) *R,R*-**1** (2.5 mM total concentration) after 1 and 3 hours, respectively.

The experiment was done to prove the similar behaviour of the NDI enantiomers *R,R*-**1** and *S,S*-**1**.

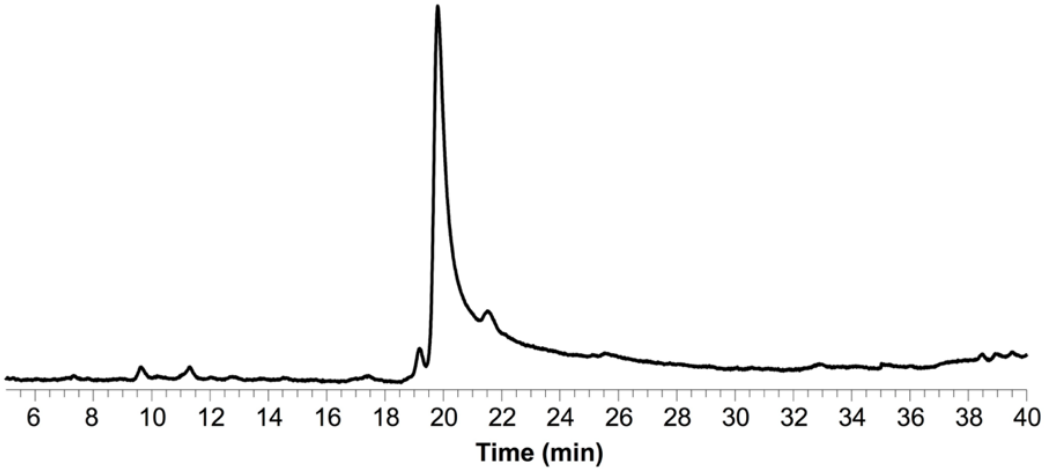


Figure 2.72. Reverse-phase HPLC analysis of *R,R*-**2** (5 mM total concentration) library. Absorbance was recorded at 389 nm.

Testing reversibility with DTT

Both libraries were reformed during the experiments with DTT, proving their reversible character, as shown in Figures 2.73 – 2.75 below. **Cat II *RSRR*** has reformed in a lower yield, because the concentration was reduced compared to **Cat I *RRRR***.

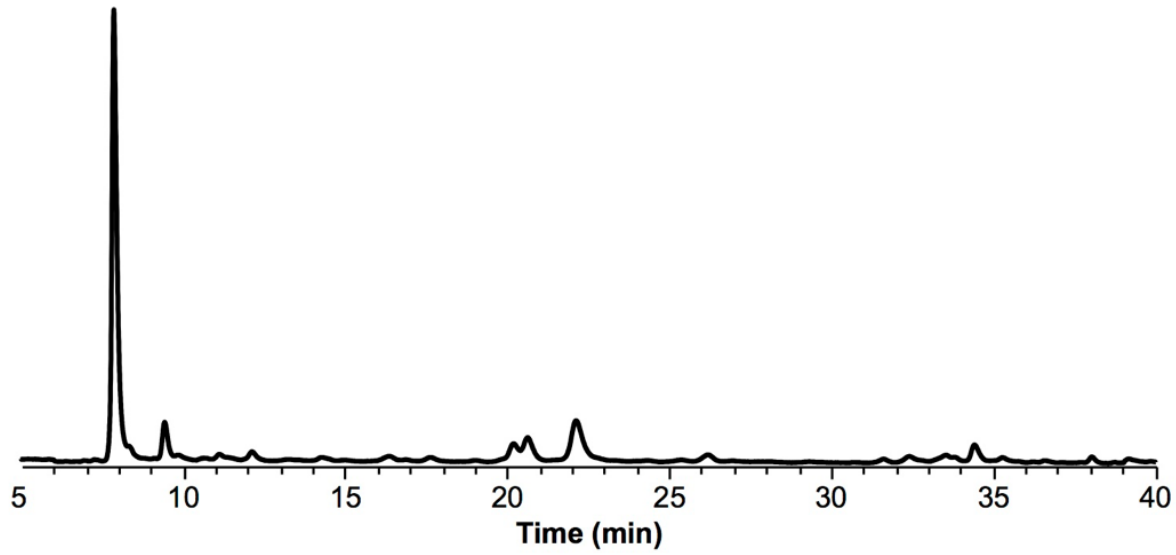


Figure 2.73. Reverse-phase HPLC analysis of **Cat I *RRRR*** isolated with DTT.

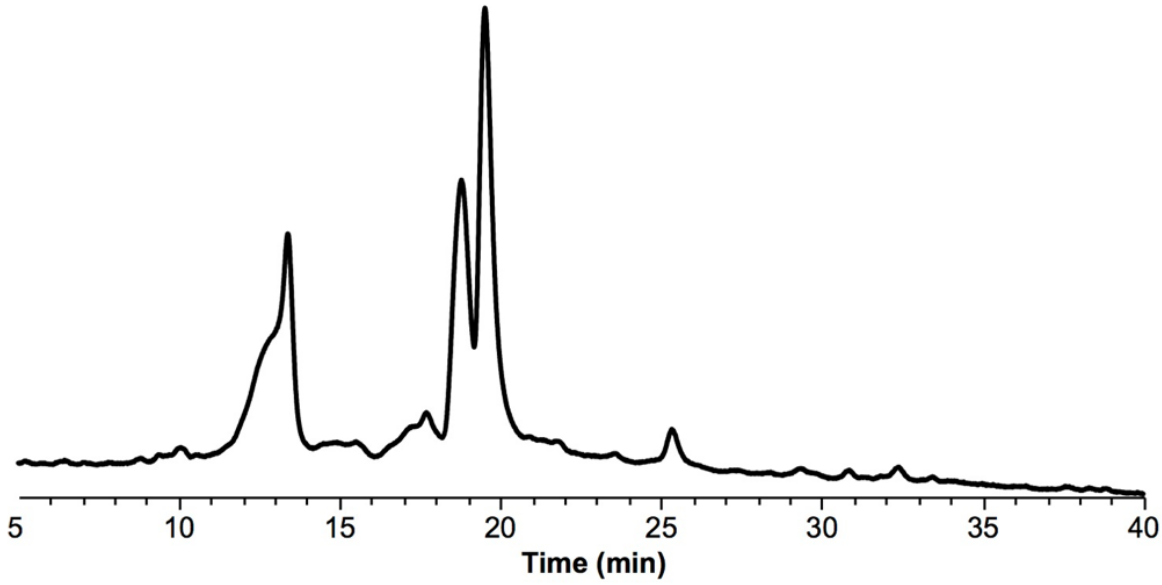


Figure 2.74. Reverse-phase HPLC analysis of **Cat II *RSRR*** isolated with DTT.

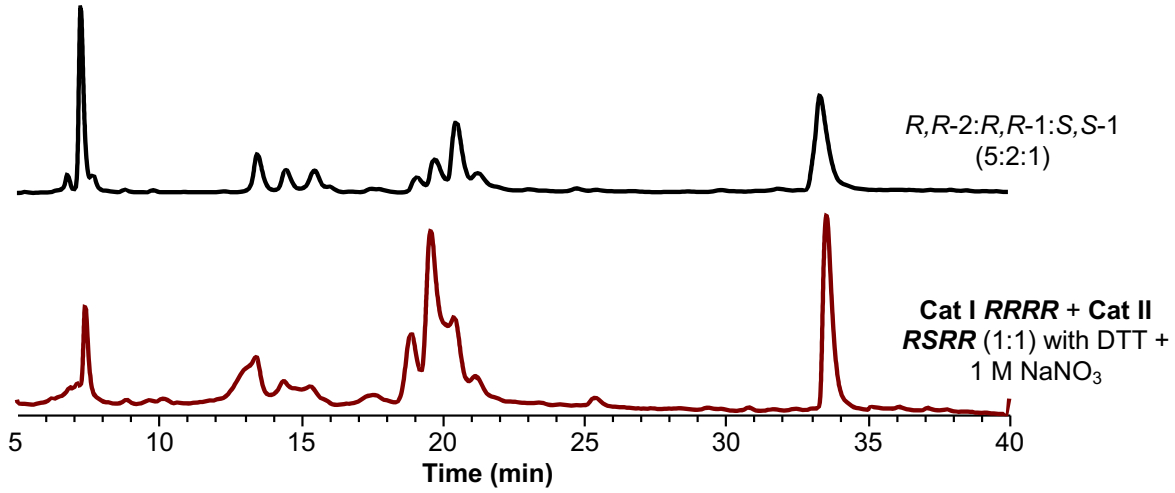


Figure 2.75. Reverse-phase HPLC analysis *R,R*-2:-*R,R*-1:*S,S*-1 (5:2:1 molar ratio, 5 mM total concentration) showed in comparison with **Cat I *RRRR*** and **Cat II *RSRR*** mixture (1:1 molar concentration) with DTT and 1 M $NaNO_3$.

2.12. DCLSim

DCLs containing the building blocks with and without the NDI-serine competitors were simulated using DCLsim.[14] The results are presented below. For each simulated DCL, we present the building block (and competitor) concentration. We also provide the Gibbs free energy of formation of each library member from its constituent building blocks (ΔG) as well the simulated and observed DCL distribution in percentages. In the DCLs containing competitors, the K_a column refers to the association constant for the competitor to the corresponding library member. We indicated that the competitor does not interact with a library member by adding a "0" in the K_a column. The equilibria responsible for the DCL distribution are presented under each table.

1. DCL of *R,R*-**1** and *R,R*-**2**:

Simulated Conditions	Conc (M)
R,R-**2**	0.0026
R,R-**1**	0.0024

Species	ΔG	Sim %	Obs %
Cat I *RRRR*	-33.66	76.71	75.83
X	-4.43	15.93	16.38
DDDA (*RRRR)* macrocycle	-28.51	7.36	7.79

2 *R,R*-**1** + 2 *R,R*-**2** ⇌ **Cat I *RRRR*** $\Delta G = -RTln\ \frac{[\mathbf{Cat\ I}\ \boldsymbol{RRRR}]}{[R,R-\mathbf{1}]^2[R,R-\mathbf{2}]^2}$

R,R-**1** + *R,R*-**2** ⇌ **X** $\Delta G = -RTln\ \frac{[\mathbf{X}]}{[R,R-\mathbf{1}]^1[R,R-\mathbf{2}]^1}$

R,R-**1** + 3 *R,R*-**2** ⇌ *DDDA* (*RRRR*) macrocycle

$\Delta G = -RTln\ \frac{[DDDA\ (RRRR)\ \text{macrocycle}]}{[R,R-\mathbf{1}]^1[R,R-\mathbf{2}]^3}$

2. DCL of *R,R*-**1** and *R,R*-**2** with NDI-*R,R*-serine:

Simulated Conditions	Conc (M)
R,R-**2**	0.00255
R,R-**1**	0.00245
NDI-*R,R*-serine	0.005

Species	ΔG	K_a (M^{-1})	K_a i.d. in Fig. S25	Sim %	Obs %
Cat I *RRRR*	-33.66	0	-	74.80	76.42
X	-4.43	81.58	K_{a3}	21.62	19.63
DDDA (*RRRR*) macrocycle	-28.51	0	-	3.58	3.95

2 *R,R*-**1** + 2 *R,R*-**2** ⇌ **Cat I *RRRR*** $\Delta G = -RTln\ \frac{[\mathbf{Cat\ I}\ \boldsymbol{RRRR}]}{[R,R-\mathbf{1}]^2[R,R-\mathbf{2}]^2}$

R,R-**1** + *R,R*-**2** ⇌ **X** $\Delta G = -RTln\ \frac{[\mathbf{X}]}{[R,R-\mathbf{1}]^1[R,R-\mathbf{2}]^1}$

R,R-**1** + 3 *R,R*-**2** ⇌ *DDDA* (*RRRR*) macrocycle

$$\Delta G = -RTln\frac{[DDDA\ (RRRR)\ \text{macrocycle}]}{[R,R-\mathbf{1}]^1[R,R-\mathbf{2}]^3}$$

X + NDI-*R*,*R*-serine ⇌ [X· NDI-*R*,*R*-serine] $K_{a3} = \frac{[\mathrm{X\cdot\ NDI}-R,R-\mathrm{serine}]]}{[\mathbf{X}]^1[\mathrm{NDI}-R,R-\mathrm{serine}]^1}$

3. DCL of *R,R*-**1** and *R,R*-**2** with NDI-*S,S*-serine:

Simulated Conditions	Conc (M)
R,R-**2**	0.00263
R,R-**1**	0.00237
NDI-*S,S*-serine	0.005

Species	ΔG	K_a (M^{-1})	K_a i.d. in Fig. S24	Sim %	Obs %
Cat I *RRRR*	-33.66	0	-	62.21	62.71
X	-4.43	233.11	K_{a2}	28.88	26.25
DDDA (*RRRR*) macrocycle	-28.51	0	-	8.91	11.04

2 *R,R*-**1** + 2 *R,R*-**2** ⇌ **Cat I *RRRR*** $\Delta G = -RTln\frac{[\mathbf{Cat\ I}\ \boldsymbol{RRRR}]}{[R,R-\mathbf{1}]^2[R,R-\mathbf{2}]^2}$

R,R-**1** + *R,R*-**2** ⇌ **X** $\Delta G = -RTln\frac{[\mathbf{X}]}{[R,R-\mathbf{1}]^1[R,R-\mathbf{2}]^1}$

R,R-**1** + 3 *R,R*-**2** ⇌ *DDDA* (*RRRR*) macrocycle

$$\Delta G = -RTln\frac{[DDDA\ (RRRR)\ \text{macrocycle}]}{[R,R-\mathbf{1}]^1[R,R-\mathbf{2}]^3}$$

X + NDI-*S,S*-serine ⇌ [X· NDI-*S,S*-serine] $K_{a2} = \frac{[\mathrm{X\cdot\ NDI}-S,S-\mathrm{serine}]]}{[\mathbf{X}]^1[\mathrm{NDI}-S,S-\mathrm{serine}]^1}$

4. DCL of *S,S*-**1** and *R,R*-**2**:

Simulated Conditions	Conc (M)
R,R-**2**	0.0025
S,S-**1**	0.0025

Species	ΔG	Sim %	Obs %
Cat II *RSRR*	-32.75	20.89	21.26
Z	-3.88	29.65	29.41
Y	-3.93	30.03	26.11
DAAA (*RSSS*) macrocycle	-32.49	17.55	18.17
DAA (*RSS*) macrocycle	0	0.01	2.89
S,S-**1** dimer	2.98	1.87	2.15

S,S-**1** + 3 *R,R*-**2** ⇌ **Cat II *RSRR*** $\Delta G = -RTln\frac{[\mathbf{Cat\ II}\ \boldsymbol{RSRR}]}{[S,S-\mathbf{1}]^1[R,R-\mathbf{2}]^3}$

S,S-**1** + *R,R*-**2** ⇌ **Z** $\Delta G = -RTln\frac{[\mathbf{Z}]}{[R,R-\mathbf{1}]^1[R,R-\mathbf{2}]^1}$

2 *R,R*-**2** ⇌ **Y** $\Delta G = -RTln\frac{[\mathbf{Y}]}{[R,R-\mathbf{2}]^2}$

3 *S,S*-**1** + *R,R*-**2** ⇌ *DAAA* (*RSSS*) macrocycle

$$\Delta G = -RTln\ \frac{[DAAA\ (RSSS)\ \text{macrocycle}]}{[S,S-\mathbf{1}]^3[R,R-\mathbf{2}]^1}$$

2 *S,S*-**1** + *R,R*-**2** ⇌ *DAA* (RSS) macrocycle

$$\Delta G = -RTln\ \frac{[DAA\ (RSS)\ \text{macrocycle}]}{[S,S-\mathbf{1}]^2[R,R-\mathbf{2}]^1}$$

2 *S,S*-**1** ⇌ *S,S*-**1** dimer $\Delta G = -RTln\ \frac{[S,S-\mathbf{1}\ \text{dimer}]^1}{[S,S-\mathbf{1}]^2}$

5. DCL of *S,S*-**1** and *R,R*-**2** with NDI-*S,S*-serine:

Simulated Conditions	Conc (M)
R,R-**2**	0.0025
S.S-**1**	0.0025
NDI-*S,S*-serine	0.005

Species	ΔG	K_a (M^{-1})	K_a i.d. in Fig. S24	Sim %	Obs %
Cat II *RSRR*	-32.75	0	-	16.20	15.91
Z	-3.88	84.9	K_{a8}	33.43	30.80
Y	-3.93	46.4	K_{a5}	29.30	27.76
DAAA (*RSSS*) macrocycle	-32.49	88.4	-	19.58	18.67
DAA (*RSS*) macrocycle	0	0	-	0.01	3.87
S,S-**1** dimer	2.98	0	-	1.48	2.97

S,S-**1** + 3 *R,R*-**2** ⇌ **Cat II *RSRR*** $\Delta G = -RTln\ \frac{[\mathbf{Cat\ II}\ \boldsymbol{RSRR}]}{[S,S-\mathbf{1}]^1[R,R-\mathbf{2}]^3}$

S,S-**1** + *R,R*-**2** ⇌ **Z** $\Delta G = -RTln\ \frac{[\mathbf{Z}]}{[R,R-\mathbf{1}]^1[R,R-\mathbf{2}]^1}$

2 *R,R*-**2** ⇌ **Y** $\Delta G = -RTln\ \frac{[\mathbf{Y}]}{[R,R-\mathbf{2}]^2}$

3 S,S-**1** + *R,R*-**2** ⇌ *DAAA* (*RSSS*) macrocycle

$$\Delta G = -RTln\ \frac{[DAAA\ (RSSS)\ \text{macrocycle}]}{[S,S-\mathbf{1}]^3[R,R-\mathbf{2}]^1}$$

2 *S,S*-**1** + *R,R*-**2** ⇌ *DAA* (*RSS*) macrocycle

$$\Delta G = -RTln\ \frac{[DAA\ (RSS)\ \text{macrocycle}]}{[S,S-\mathbf{1}]^2[R,R-\mathbf{2}]^1}$$

2 *S,S*-**1** ⇌ *S,S*-**1** dimer $\Delta G = -RTln\ \frac{[S,S-\mathbf{1}\ \text{dimer}]^1}{[S,S-\mathbf{1}]^2}$

Z + NDI-*S,S*-serine ⇌ [Z · NDI-*S,S*-serine] $K_{a8} = \frac{[\text{Z} \cdot \text{NDI}-S,S-\text{serine}]}{[\mathbf{Z}]^1[\text{NDI}-S,S-\text{serine}]^1}$

Y + NDI-*S,S*-serine ⇌ [Y · NDI-*S,S*-serine] $K_{a5} = \frac{[\text{Y} \cdot \text{NDI}-S,S-\text{serine}]}{[\mathbf{Y}]^1[\text{NDI}-S,S-\text{serine}]^1}$

DAAA (*RSSS*) macrocycle + NDI-*S,S*-serine ⇌ [*DAAA* (*RSSS*) macrocycle · NDI-*S,S*-serine]

$$K_a = \frac{[DAAA\ (RSSS)\ \text{macrocyle} \cdot \text{NDI}-S,S-\text{serine}]}{[\boldsymbol{D}AAA\ (RSSS)\ \text{macrocyle}]^1[\text{NDI}-S,S-\text{serine}]^1}$$

6. DCL of *S,S*-**1** and *R,R*-**2** with NDI-*R,R*-serine:

Simulated Conditions	**Conc (M)**
R,R-**2**	0.0025
S.S-**1**	0.0025
NDI-*R,R*-serine	0.005

Species	ΔG	K_a (M^{-1})	K_a i.d. in Fig. S24	Sim %	Obs %
Cat II *RSRR*	-32.75	0	-	8.37	7.76
Z	-3.88	285.3	K_{a9}	31.59	30.28
Y	-3.93	297	K_{a6}	32.73	29.24
DAAA (*RSSS*) macrocycle	-32.49	566.7	-	26.48	24.76
DAA (*RSS*) macrocycle	0	0	-	0.01	5.35
S,S-**1** dimer	2.98	0	-	0.83	2.61

S,S-**1** + 3 *R,R*-**2** ⇌ **Cat II *RSRR*** $\Delta G = -RTln\frac{[\textbf{Cat II } \boldsymbol{RSRR}]}{[S,S-\mathbf{1}]^1[R,R-\mathbf{2}]^3}$

S,S-**1** + *R,R*-**2** ⇌ **Z** $\Delta G = -RTln\frac{[\mathbf{Z}]}{[R,R-\mathbf{1}]^1[R,R-\mathbf{2}]^1}$

2 R,R-**2** ⇌ **Y** $\Delta G = -RTln\frac{[\mathbf{Y}]}{[R,R-\mathbf{2}]^2}$

3 S,S-**1** + *R,R*-**2** ⇌ *DAAA* (*RSSS*) macrocycle

$$\Delta G = -RTln\frac{[DAAA\ (RSSS)\ \text{macrocycle}]}{[S,S-\mathbf{1}]^3[R,R-\mathbf{2}]^1}$$

2 *S,S*-**1** + *R,R*-**2** ⇌ *DAA* (*RSS*) macrocycle

$$\Delta G = -RTln\frac{[DAA\ (RSS)\ \text{macrocycle}]}{[S,S-\mathbf{1}]^2[R,R-\mathbf{2}]^1}$$

2 *S,S*-**1** ⇌ *S,S*-**1** dimer $\Delta G = -RTln\frac{[S,S-\mathbf{1}\ \text{dimer}]^1}{[S,S-\mathbf{1}]^2}$

Z + NDI-*R,R*-serine ⇌ [Z · NDI-*R,R*-serine] $K_{a9} = \frac{[\mathrm{Z}\cdot \mathrm{NDI}-R,R-\mathrm{serine}]}{[\mathbf{Z}]^1[\mathrm{NDI}-R,R-\mathrm{serine}]^1}$

Y + NDI-*R,R*-serine ⇌ [Y · NDI-*R,R*-serine] $K_{a6} = \frac{[\mathrm{Y}\cdot \mathrm{NDI}-R,R-\mathrm{serine}]}{[\mathbf{Y}]^1[\mathrm{NDI}-R,R-\mathrm{serine}]^1}$

DAAA (*RSSS*) macrocycle + NDI-*R,R*-serine ⇌ [*DAAA (RSSS)* macrocycle · NDI-*R,R*-serine]

$$K_a = \frac{[DAAA\ (RSSS)\ \text{macrocyle}\cdot \mathrm{NDI}-R,R-\mathrm{serine}]}{[DAAA\ (RSSS)\ \text{macrocyle}]^1[\mathrm{NDI}-R,R-\mathrm{serine}]^1}$$

2.13. References

1 M. D. Hanwell, D. E. Curtis, D. C. Lonie, T. Vandermeersch, E. Zurek and G. R. Hutchison, *J. Cheminformatics*, 2012, **4**, 17.

2 I. Stolić, K. Molčanov, G. Kovačević, B. Kojić-Prodić and M. Bajić, *Struct. Chem.*, 2012, **23**, 425–432.

3 R. J. Abraham, N. Aboitiz, M. Filippi, E. Genesio, P. Piaggio and F. Sancassan, *Magn. Reson. Chem.*, 2015, **53**, 498–508.

4 J. F. Galan, J. Brown, J. L. Wildin, Z. Liu, D. Liu, G. Moyna and V. Pophristic, *J. Phys. Chem. B*, 2009, **113**, 12809–12815.

5 B. Kuhn, P. Mohr and M. Stahl, *J. Med. Chem.*, 2010, **53**, 2601–2611.

6 M. Parvez, P. C. Unangst, D. T. Connor and M. D. Mullican, *Acta Crystallogr. C*, 1991, **47**, 611–613.

7 P. Pengo, G. D. Pantoş, S. Otto and J. K. M. Sanders, *J. Org. Chem.*, 2006, **71**, 7063–7066.

8 H. Y. Au-Yeung, G. D. Pantoş and J. K. M. Sanders, *Proc. Natl. Acad. Sci.*, 2009, **106**, 10466–10470.

9 H. Y. Au-Yeung, F. B. L. Cougnon, S. Otto, G. D. Pantoş and J. K. M. Sanders, *Chem. Sci.*, 2010, **1**, 567–574.

10 S. Alem, S. Wakim, J. Lu, G. Robertson, J. Ding and Y. Tao, *ACS Appl. Mater. Interfaces*, 2012, **4**, 2993–2998.

11 M. L. Keshtov, S. A. Kuklin, A. R. Khokhlov, I. O. Konstantinov, N. V. Nekrasova, Z. Xie, S. Biswas and G. D. Sharma, *New J. Chem.*, 2018, **42**, 1626–1633.

12 L. Liu, F. Li, C. Wu, T. Tang, X. Cheng, X. Luo, J. Li, *Macromol. Res.*, 2018, **27**, 534-542.

13 J. Du, A. Fortney, K. E. Washington, C. Bulumulla, P. Huang, D. Dissanayake, M. C. Biewer, T. Kowalewski and M. C. Stefan, *ACS Appl. Mater. Interfaces*, 2016, **8**, 33025–33033.

14 P. T. Corbett, S. Otto and J. K. M. Sanders, *Chem. - Eur. J.*, 2004, **10**, 3139–3143.

15 H. Y. Au-Yeung, G. D. Pantoş and J. K. M. Sanders, *Angew. Chem. Int. Ed.*, 2010, **49**, 5331–5334.

16 H. Y. Au-Yeung, G. D. Pantoş and J. K. M. Sanders, *J. Org. Chem.*, 2011, **76**, 1257–1268.

17 H. Y. Au-Yeung, G. D. Pantoş and J. K. M. Sanders, *J. Am. Chem. Soc.*, 2009, **131**, 16030–16032.

Chapter 3. Discovery of an all-donor catenane

Catenanes have become a curiosity since their advent about half a century ago.[1–3] Many groups have invested a lot of effort trying to obtain [2]catenanes in high yields. Nowadays, the synthesis of [2]catenane is considered almost trivial, and the aim is to make high-order catenanes (*e.g.,* [n]catenanes, $n \geq 3$) or finding applications for these structures. Probably the most common type of catenanes is the one in which four aromatic units are held together by non-covalent π-π interactions. Among these, the studies reported in the literature mostly focus on donor-acceptor-based catenanes. Back in the days when research on this topic was started, it was thought that the only possible stacking is a *DADA* type (*D* = donor, *A* = acceptor), due to favourable interaction between a donor and an acceptor. However, this general belief has come to an end when Sanders' and Pantoș' groups have showed that other stacking sequences are also possible.[4–6] So far, various stacking arrangements were synthesised, including: *DADD*, *DADA*, *DAAD*, *AADA* and *AAAA*. The *AAAA* stacking (all-acceptor [2]catenane) was independently obtained by different groups. One of them is Sanders' group; in this case an alkyl chain links two NDI aromatic cores.[6] The second one belongs to Stoddart's group, which followed the threading and clipping approach).[7] Another all-acceptor [2]catenane consisted of four perylenediimide (PDI) cores, which was assembled in an organic solvent (CH_2Cl_2).[8] However, there is a missing structure in this set: a [2]catenane with a *DDDD* stacking sequence, thus an all-donor catenane (Figure 3.1a). Up to date, it was believed that this arrangement is impossible because of the π-rich nature of the components, which leads to repulsions between the aromatics. The only example of an all-donor [2]catenane comes from Stoddart's group, yet this is an unusual example. In this case, one of the components in the stacking is not aromatic, but it contains an extended conjugated system, which is a π-donor molecule (Figure 3.1b).[9]

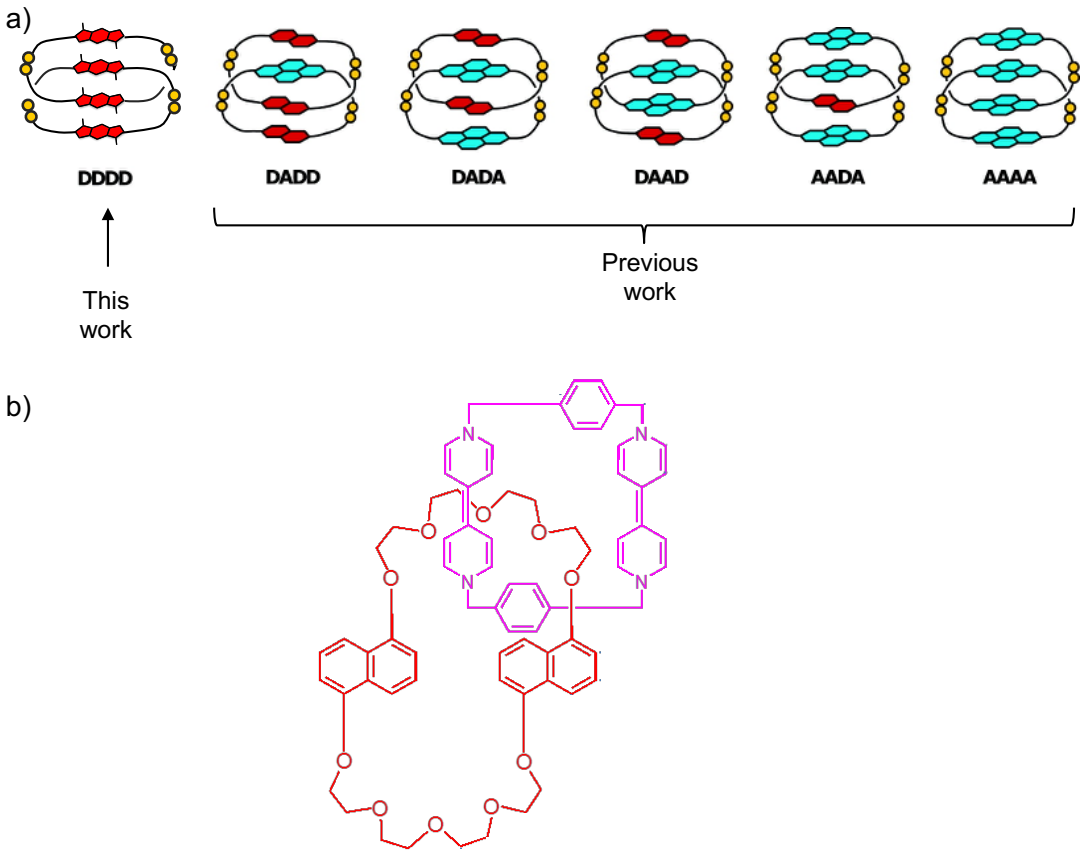


Figure 3.1. a) Graphical representations of the new stacking reported in this work alongside with the other types of catenanes published before and b) the previously synthesised all-donor catenane by Stoddart's group.[9]

3.1. Analysis of *R,R*-2 library

One of the molecules that was introduced in Chapter 2 is a new aromatic π-donor building block, BDT, which consists of a benzo[1,2-*b*:4,5-*b*']dithiophene (*R,R*-**2** or *S,S*-**2**; BDT) with *R* or *S* cysteine appendages for disulphide exchange. This can be synthesised in seven steps starting from 1,4-dibromo-2,5-terephthalic acid. The next discussion will focus on the behaviour of *R,R*-**2** and *S,S*-**2** in enantio-pure DCLs and a racemic mixture DCL. We will also introduce the DCC behaviour of iso-BDT (an isomer of *R,R*-**2**, a benzo[1,2-*b*:5,4-*b*']dithiophene derivative, also synthesised by us). DCLs were set up by dissolving the building block (*R,R*-**2**) in water at basic pH (8) and allow to oxidise under atmospheric conditions and the library was stirred for about three days before analysis. In the case of the BDT DCL, HPLC and LC-MS studies

revealed the formation of a homodimer (Figure 2.50 for its characterisation and Figure 2.72 for the HPLC trace). The increase of the ionic strength of the system by addition of 1 M $NaNO_3$, leads to a new peak in the chromatogram. Initial analysis of the new peak (Figure 3.2) along with its shorter retention in comparison with the homodimer suggest that this new peak would correspond to a [2]catenane (formed in 15% yield in DCL). The MS and zoom scan analyses of this new peak indicate a species containing four BDT units, while the MS/MS analysis confirms the catenated structure of the molecule (*i.e.* no fragment with *m/z* more than a dimer; Figure 3.3). The catenane was isolated for further characterisation (even if it contains a small quantity of **Y** (BDT dimer) – Figure 3.4).

3.1.1. Analysis of *DDDD* [2]catenane *RRRR*

^{1}H NMR analysis of the equilibrated library reveals three peaks in the aromatic region. One of them corresponds to the homodimer, while the presence of the other two, more shielded than the one corresponding to the homodimer, also support the proposed structure of an all-donor catenane. The most shielded peak belongs to the inner benzodithiophene protons, whereas the outer one is less influenced by the ring current effect, being less shielded. The isolated catenane was subjected to 2D NOESY analysis to understand the environment of the molecule (Figure 3.5).

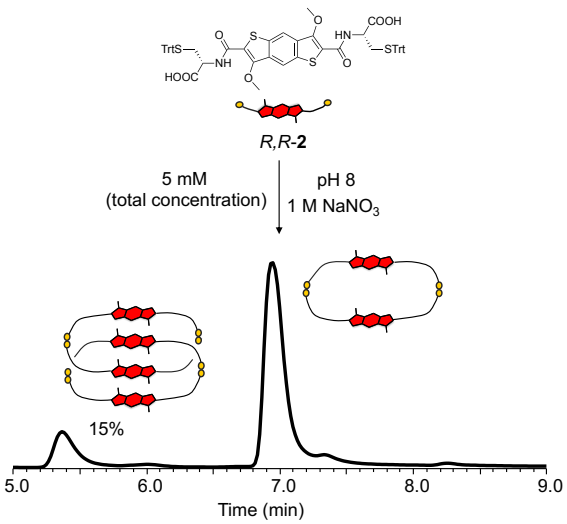


Figure 3.2. Reverse-phase HPLC analysis of *R,R*-**2** (5 mM concentration) library in the presence of 1 M $NaNO_3$ showing the formation of an all-donor catenane in 15% yield. Absorbance was recorded at 391 nm.

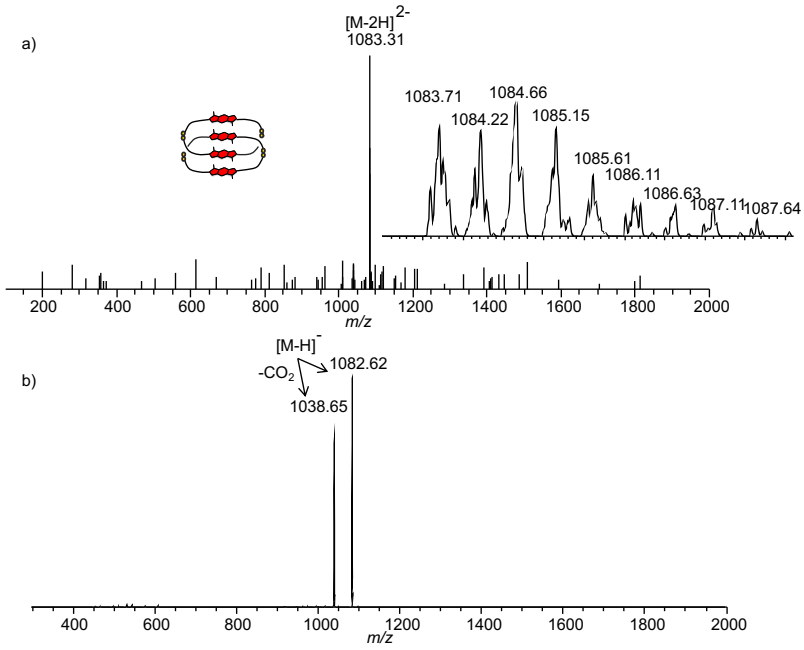


Figure 3.3. a) MS (-*ve*) of ***DDDD* [2]catenane *RRRR***; zoom of the molecular ion is shown as an inset and b) MS/MS (-*ve*) of ***DDDD* [2]catenane *RRRR***.

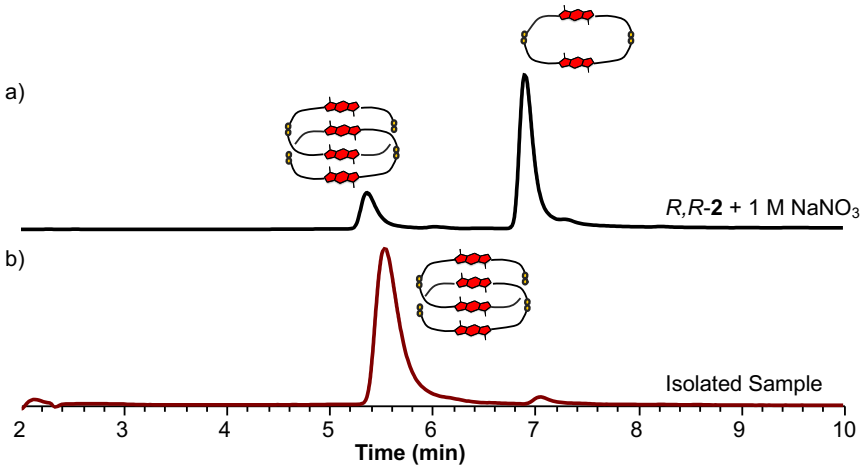


Figure 3.4. a) Reverse-phase HPLC analysis of *R,R*-**2** (5 mM concentration) library in the presence of 1 M $NaNO_3$ and b) the reverse-phase HPLC analysis of the isolated ***DDDD* [2]catenane *RRRR***. Absorbance was recorded at 391 nm.

The NOESY spectrum shows cross peaks between the protons in aromatic region, suggesting a slightly offset parallel arrangement of the aromatic cores. The cross peaks between the methoxy groups and the aromatic protons indicate their close proximity (Figure 3.6). On this basis, this new species was assigned as the first ever all-donor catenane obtained through DCC. Attempts to enhance the yield of this

[2]catenane were performed by increasing the salt content ($NaNO_3$) to 5 and 10 M. Unfortunately, these attempts failed, and the maximum yield was 15% (Figure 3.7).

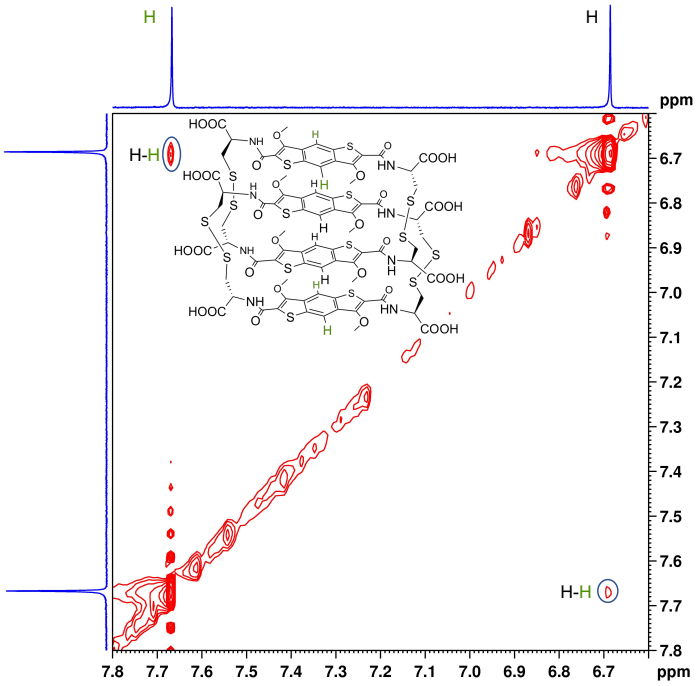


Figure 3.5. 2D NOESY (red) spectrum (500 MHz, 298 K) of ***DDDD*** **[2]catenane** ***RRRR*** in the aromatic region. The residual solvent (H_2O) was referenced at 4.79 ppm.

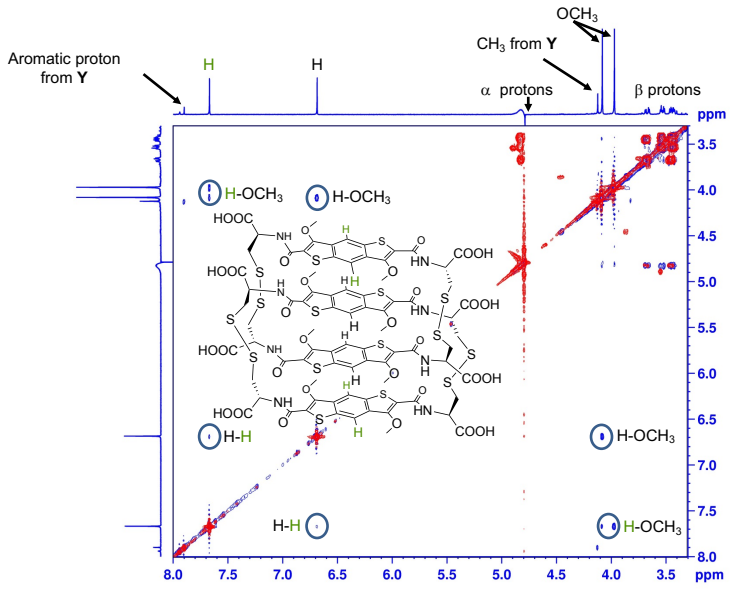


Figure 3.6. Full 2D NOESY (blue) and ^1H-^1H COSY (red) spectra (500 MHz, 298 K) of ***DDDD*** **[2]catenane** ***RRRR***. The residual solvent (H_2O) was referenced at 4.79 ppm.

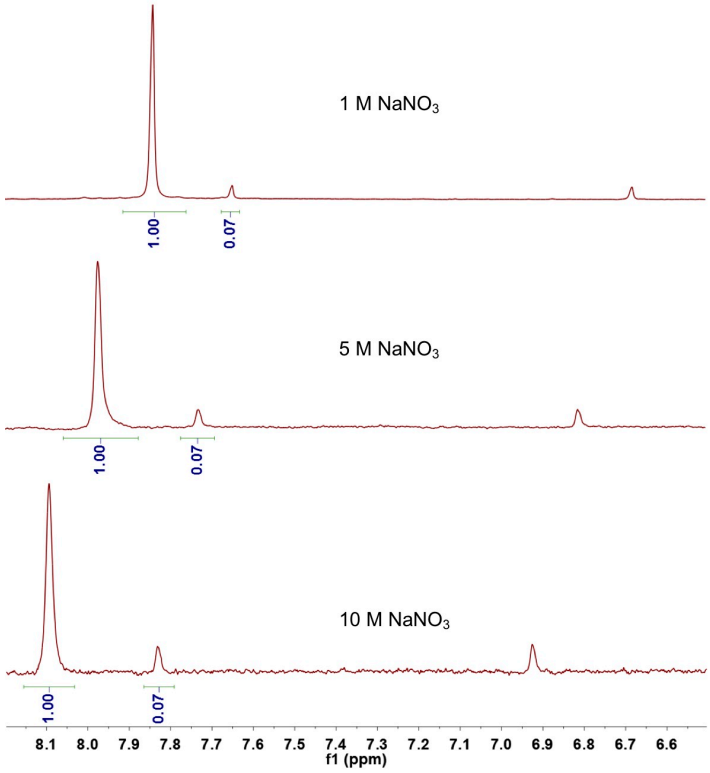


Figure 3.7. The ^{1}H NMR spectra of *R,R*-**2** (5 mM concentration) libraries in the presence of 1 M $NaNO_3$, 5 M $NaNO_3$ and 10 M $NaNO_3$.

As an all-donor catenane was thought to be impossible to form so far, we provide reasoning behind the formation of our ***DDDD*** **[2]catenane** ***RRRR***. As discussed in Chapter 2, BDT is a quasi-fused pentacyclic core. The large hydrophobic core of BDT is responsible for the formation of the all-donor catenane. In 1 M $NaNO_3$ solution, the hydrophobic effect overcomes the electronic repulsions between the electron-rich π surfaces, thus allowing the *DDDD* stack to form.

3.2. Analysis of *S,S*-2 library

We have also synthesised the other enantiomer of the building block mentioned above (*R,R*-**2**) – *S,S*-**2.** As expected, it has a similar behaviour, so the all-donor catenane, ***DDDD*** **[2]catenane** ***SSSS***, is also formed in the presence of 1 M $NaNO_3$ (Figures 3.8 – 3.13).

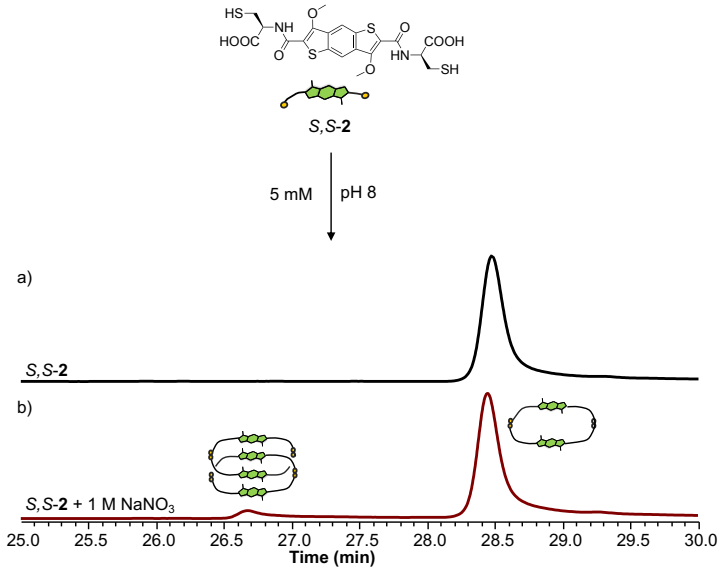


Figure 3.8. Reverse-phase HPLC analysis of *S,S*-**2** (5 mM concentration) library a) without salt and b) in the presence of 1 M $NaNO_3$. Absorbance was recorded at 391 nm.

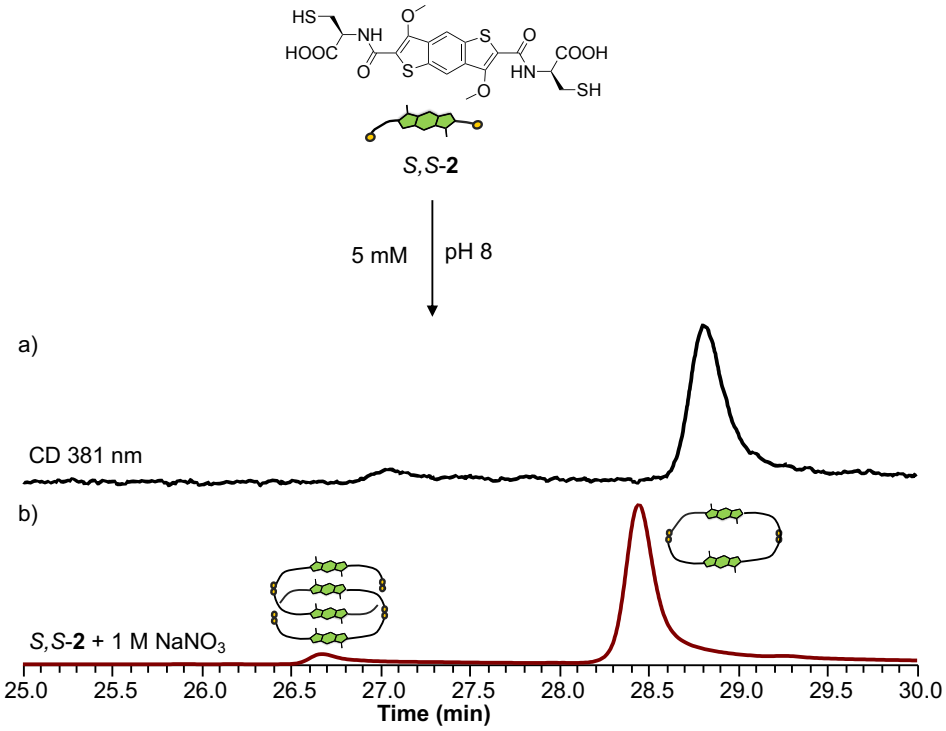


Figure 3.9. Reverse-phase HPLC analysis of *S,S*-**2** (5 mM concentration) library in the presence of 1 M $NaNO_3$ showing a) CD at 381 nm and b) absorbance at 391 nm.

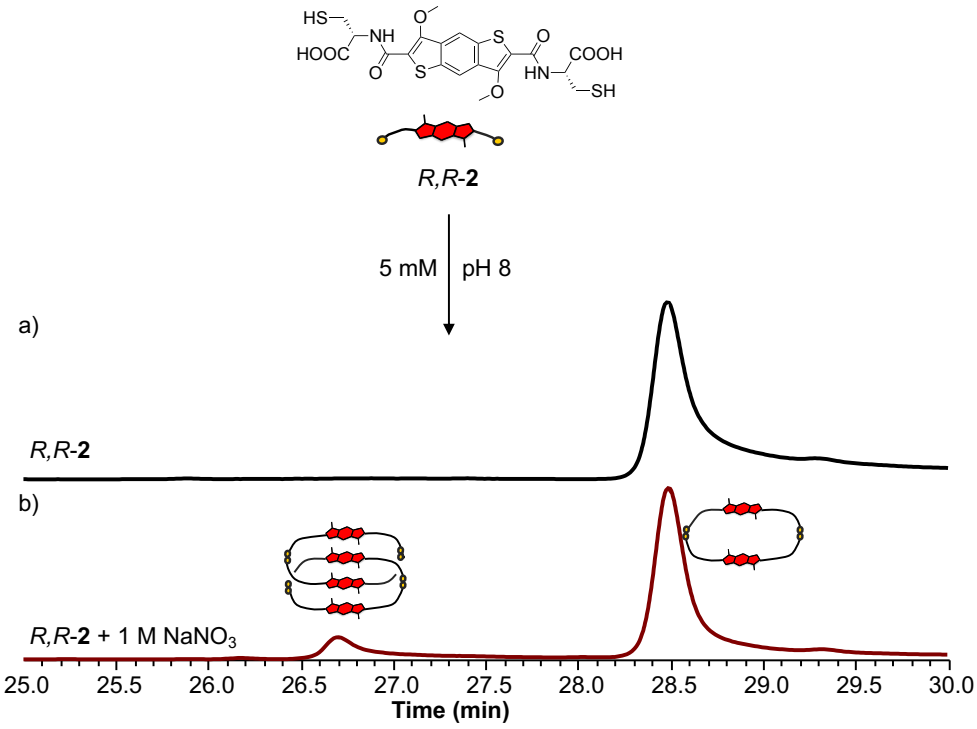


Figure 3.10. Reverse-phase HPLC analysis of *R,R*-**2** (5 mM concentration) library a) without salt and b) in the presence of 1 M $NaNO_3$. Absorbance was recorded at 391 nm.

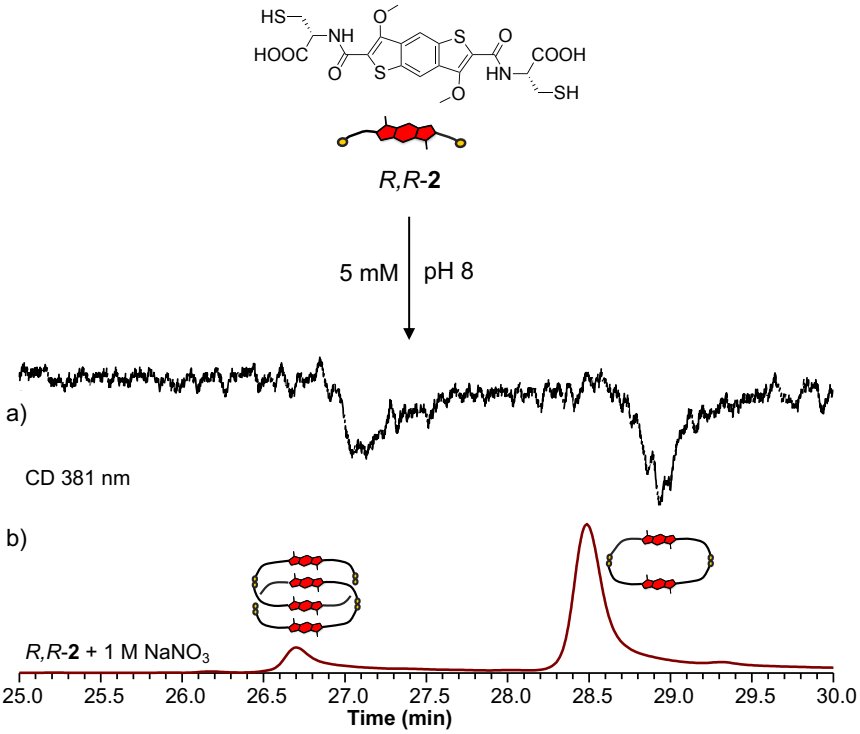


Figure 3.11. Reverse-phase HPLC analysis of *R,R*-**2** (5 mM concentration) library in the presence of 1 M $NaNO_3$ showing a) CD at 381 nm and b) absorbance at 391 nm.

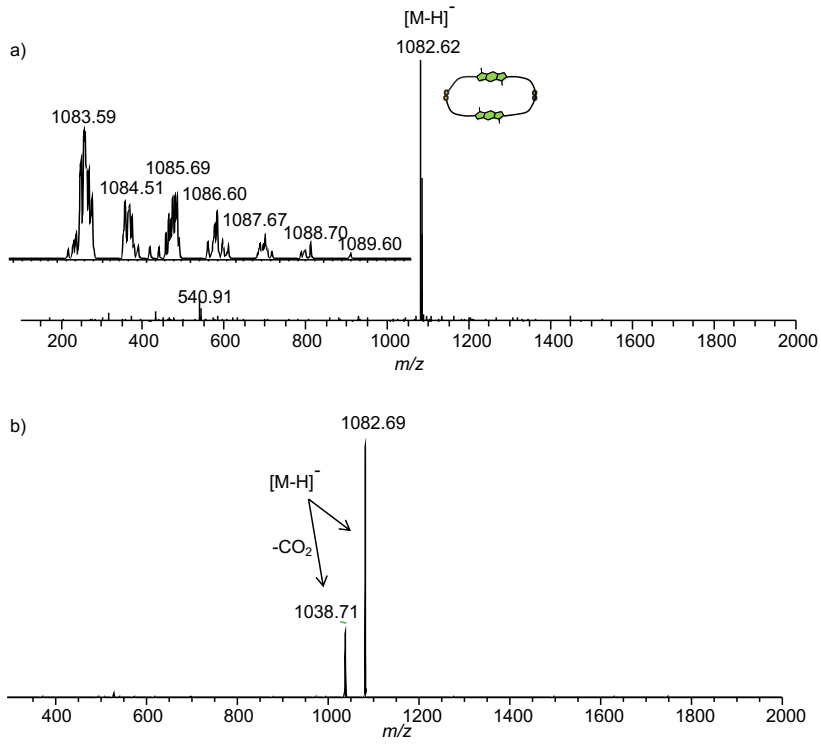


Figure 3.12. a) MS (*-ve*) of **Y2**; zoom of the molecular ion is shown as an inset and b) MS/MS (*-ve*) of **Y2**.

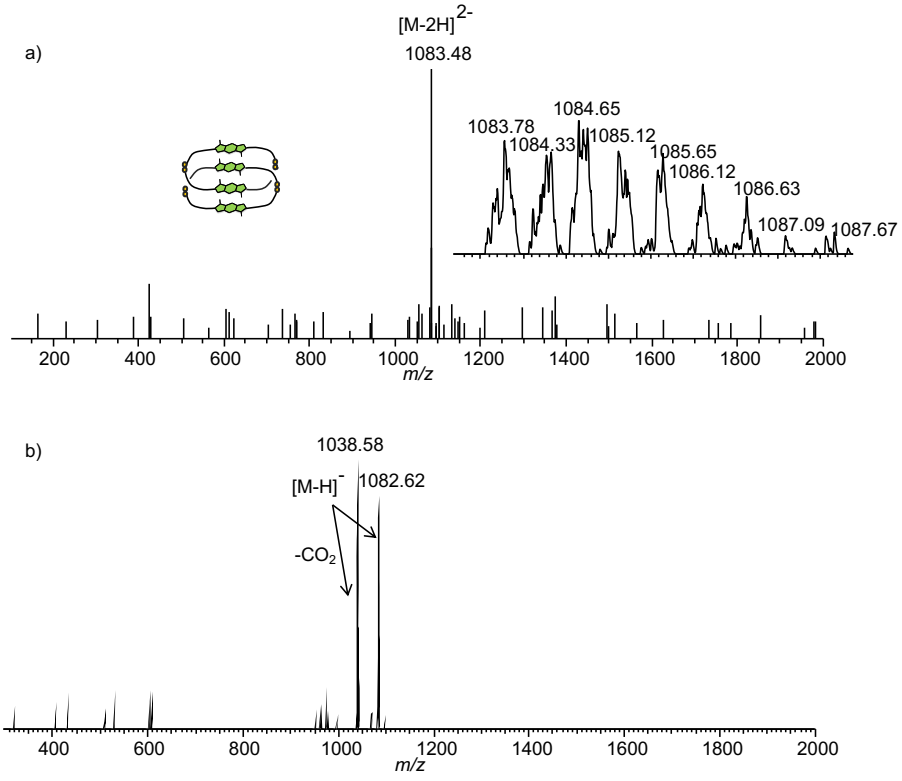
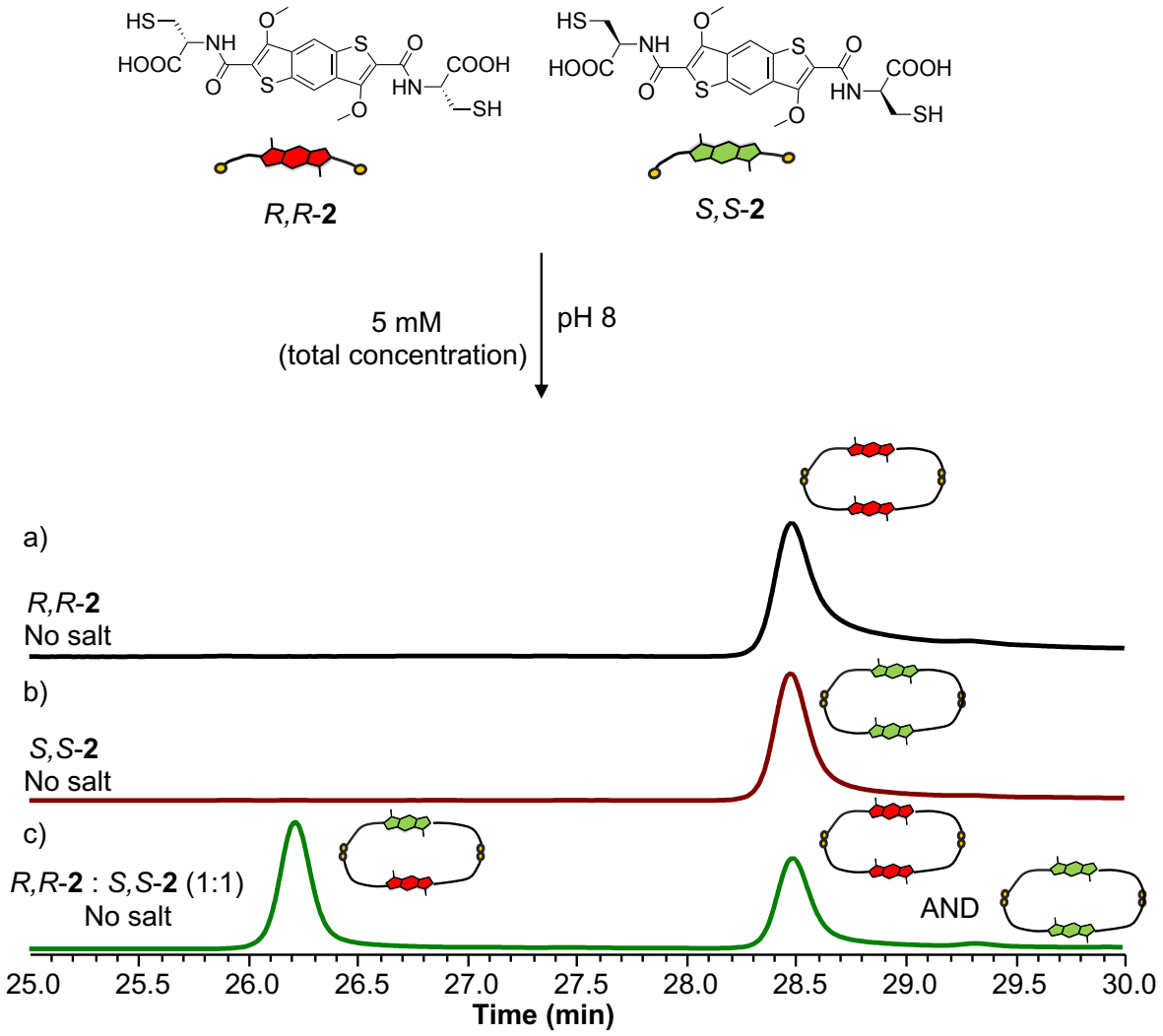


Figure 3.13. a) MS (*-ve*) of ***DDDD*** **[2]catenane** ***SSSS***; zoom of the molecular ion is shown as an inset and b) MS/MS *(-ve*) of ***DDDD*** **[2]catenane** ***SSSS***.

3.3. Analysis of *R,R*-2 and *S,S*-2 libraries

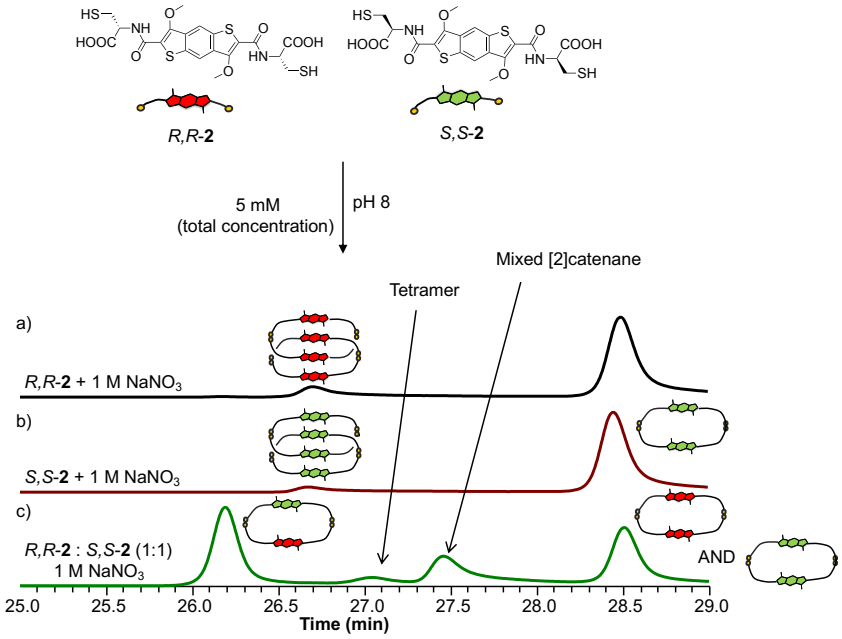
We have also set up a racemic library of *R,R*-**2** and *S,S*-**2**. In the absence of salt, three species were formed at equilibrium: **Y** (*R,R*-**2** homodimer), **Y1** (the heterochiral homodimer) and **Y2** (*S,S*-**2** homodimer) shown in Figure 3.14. When salt was added, the equilibrated library contains a heterodimer, homodimers, a heterotetramer, and a heterochiral all-donor [2]catenane. The HPLC set-up allows for single wavelength CD detection, which has provided details about the homodimer composition. If the peak were composed of equal amounts of enantiopure homodimers, then no CD signal should be observed. In our case, a positive Cotton effect has suggested the formation of **Y2** in greater amount than **Y**. It is essential to note that no homodimer catenane was formed, but rather a tetramer and heterochiral catenane. In this case, this mismatch in chirality drives the distribution of the DCL, leading to a tetramer formation rather than a catenated species (Figures 3.14 – 3.19). This might be because of a cavity with a slightly different shape than the homodimer, which disfavours threading. The space-filling representations of simulated dimers (homo and heterodimer) are in accordance with the assumption that the heterodimer has a different cavity (Figure 3.20).

We have also studied the chiroptical and UV-Vis properties of the *R*- and *S*-homodimers and the all-donor catenane (***RRRR*** version). The absorbance profile of the homodimers and the catenane look similar with minor differences. In the case of the catenane, the peaks are blue-shifted and broad due to chromophore combination (Figure 3.21). As for CD, the catenane displays a low Cotton effect on all the absorbances and different CD profile due to chromophore combination (Figure 3.22). Taking advantage of fluorescent benzodithiophene core and the lack of any fluorescence quenchers (*e.g.,* naphthalenediimide core), we also introduce the first reported fluorescent catenane obtained through a DCC approach. The quantum yield (QY) for the homodimer (**Y2**) is 0.007, whereas the QY for the catenane (*RRRR*) is about five times greater – 0.033. This shows that combined benzodithiophene cores induce an additive effect on the fluorescence outcome (Figure 3.23).



Species			
i.d.	**Y**	**Y1**	**Y2**

Figure 3.14. a) Reverse-phase HPLC analysis of *R,R*-**2** (5 mM concentration), b) *S,S*-**2** (5 mM concentration) and c) *R,R*-**2**:*S,S*-**2** (1:1 molar ratio, 5 mM concentration) libraries without salt. Absorbance was recorded at 391 nm.



Species					
i.d.	**Y**	**Y1**	**Y2**	***DDDD* [2]catenane *RRRR***	***DDDD* [2]catenane SSSS**

Figure 3.15. a) Reverse-phase HPLC analysis of *R,R*-**2** (5 mM concentration), b) *S,S*-**2** (5 mM concentration) and c) *R,R*-**2**:*S,S*-**2** (1:1 molar ratio, 5 mM concentration) libraries in the presence of 1 M $NaNO_3$. Absorbance was recorded at 391 nm.

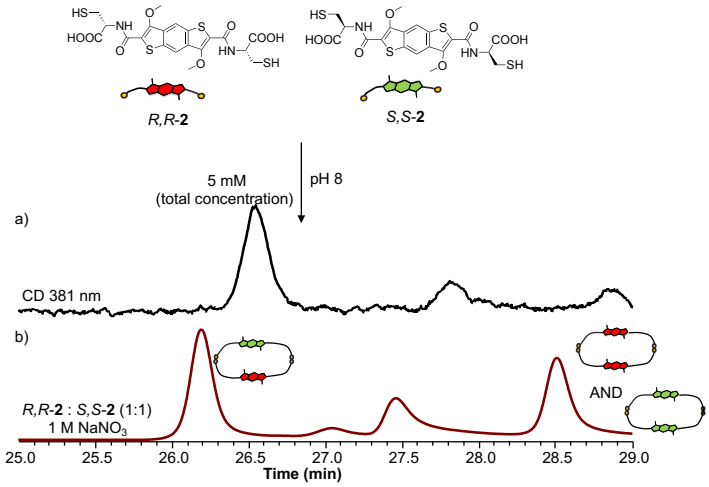


Figure 3.16. Reverse-phase HPLC analysis of *R,R*-**2**:*S,S*-**2** (1:1 molar ratio, 5 mM concentration) library in the presence of 1 M $NaNO_3$ showing a) CD at 381 nm and b) absorbance at 381 nm.

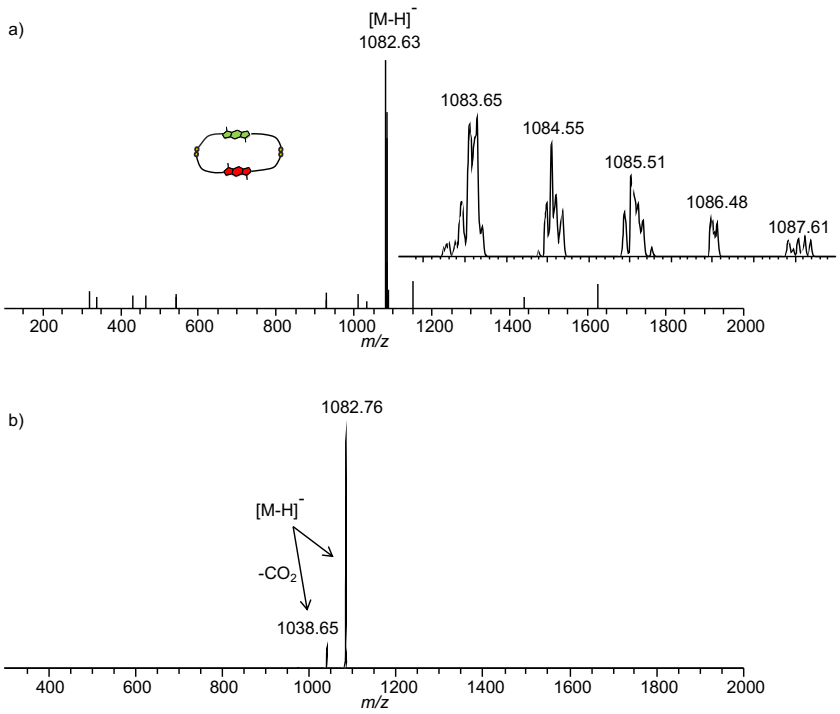


Figure 3.17. a) MS (-*ve*) of **Y1**; zoom of the molecular ion is shown as an inset and b) MS/MS (-*ve*) of **Y1**.

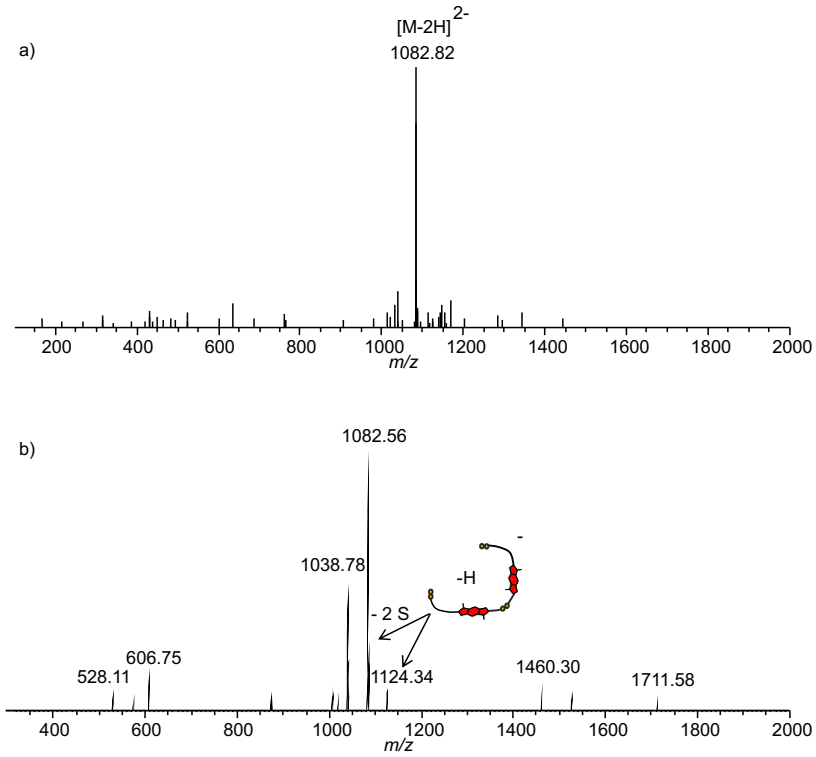


Figure 3.18. a) MS (-*ve*) of the tetramer; zoom of the molecular ion is shown as an inset and b) MS/MS (-*ve*) of the tetramer. The species with *m/z* more than a dimer reinforces the idea of a macrocycle rather than a catenane.

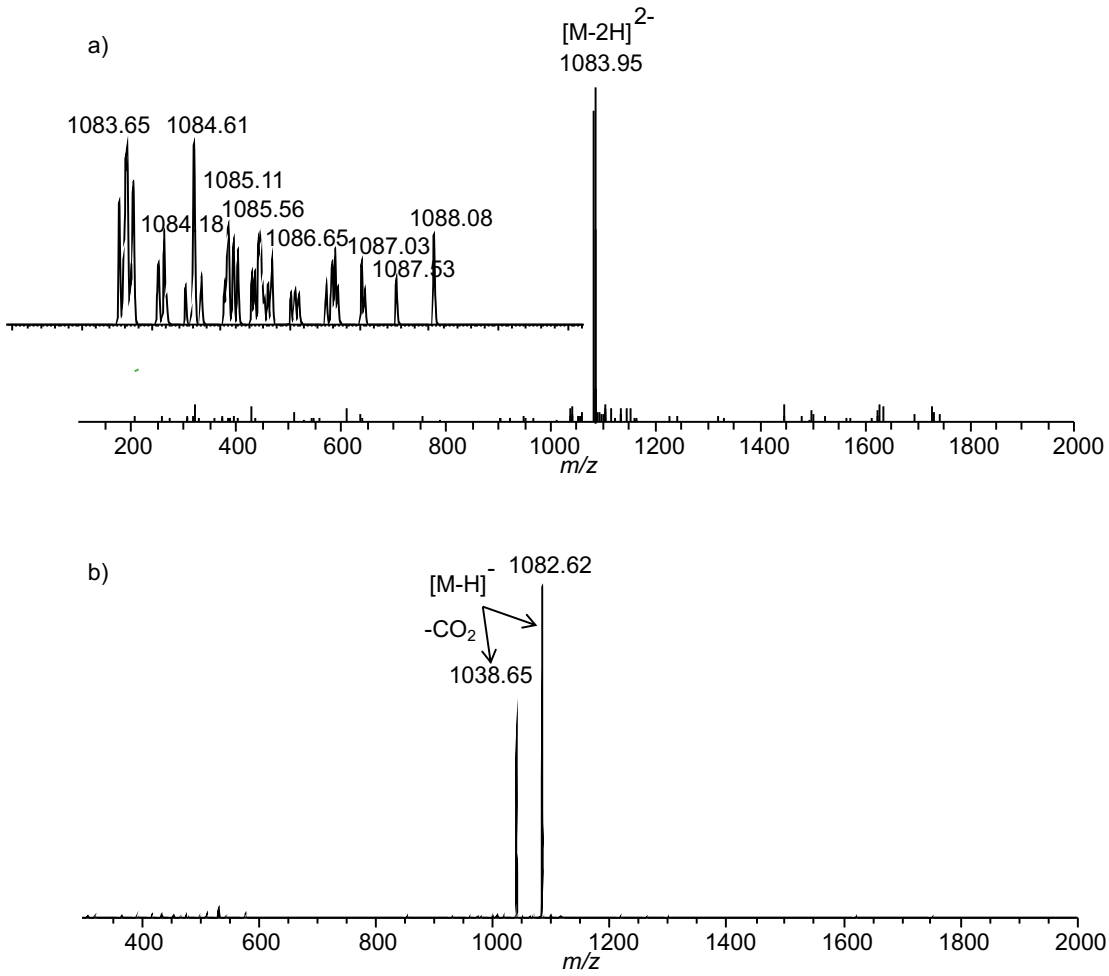


Figure 3.19. a) MS (-*ve*) of the [2]catenane with mixed chiralities; zoom of the molecular ion is shown as an inset and b) MS/MS (-*ve*) of the molecule.

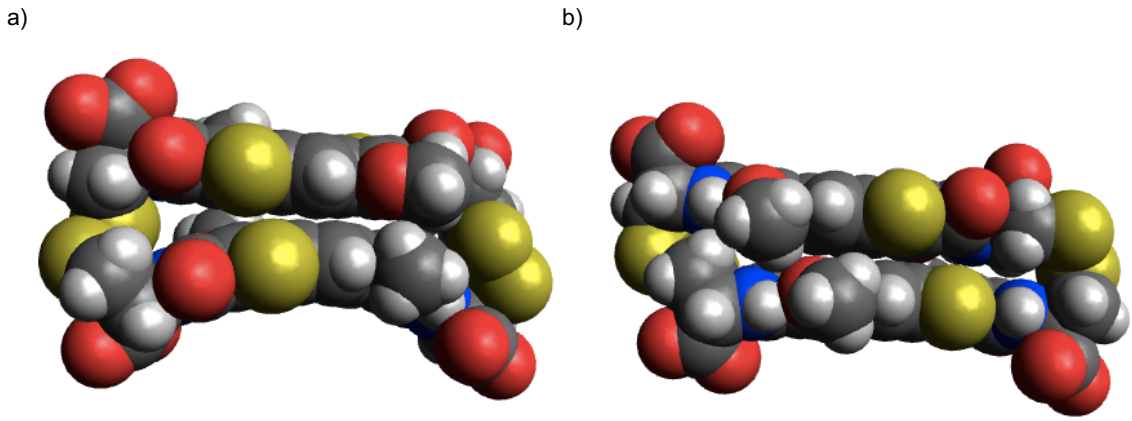
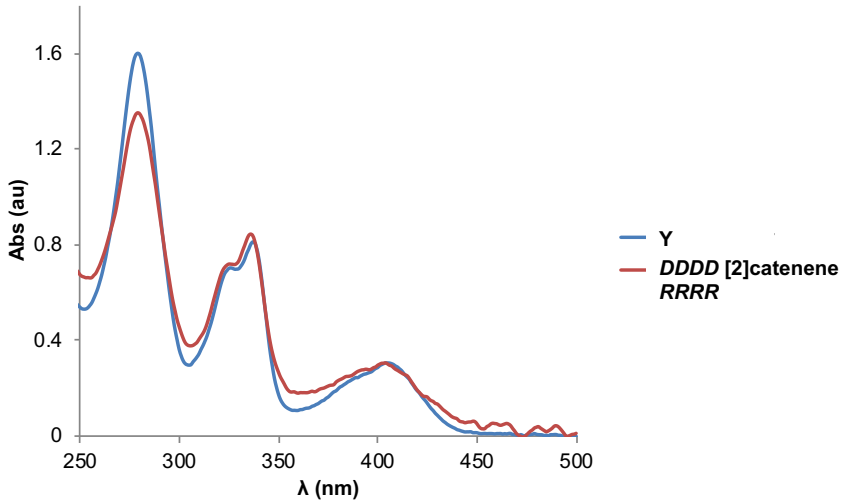
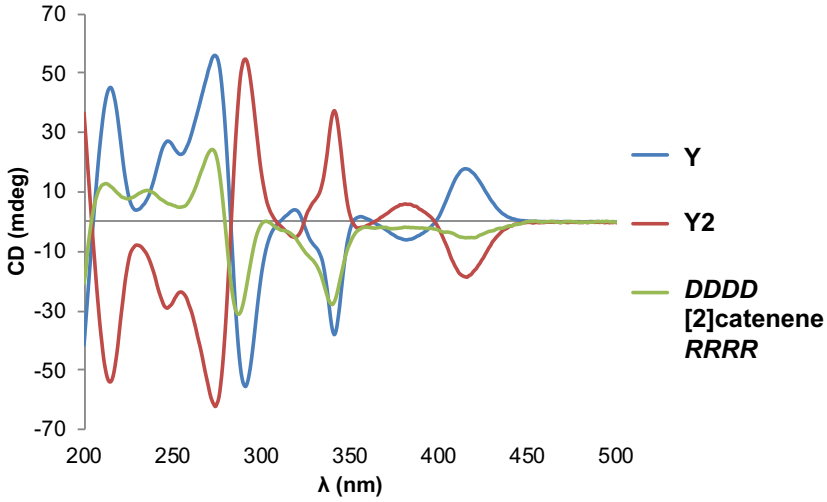


Figure 3.20. Space-filling representations of the simulated a) **Y** and b) **Y1**.



Species		
i.d.	***DDDD* [2]catenane *RRRR***	**Y**

Figure 3.21. Absorbance spectra in water of **Y** and ***DDDD* [2]catenane *RRRR*** normalised by absorbance.



Species			
i.d.	***DDDD* [2]catenane *RRRR***	**Y**	**Y2**

Figure 3.22. Circular Dichroism spectra in water of **Y**, **Y2** and ***DDDD* [2]catenane *RRRR*** normalised by absorbance.

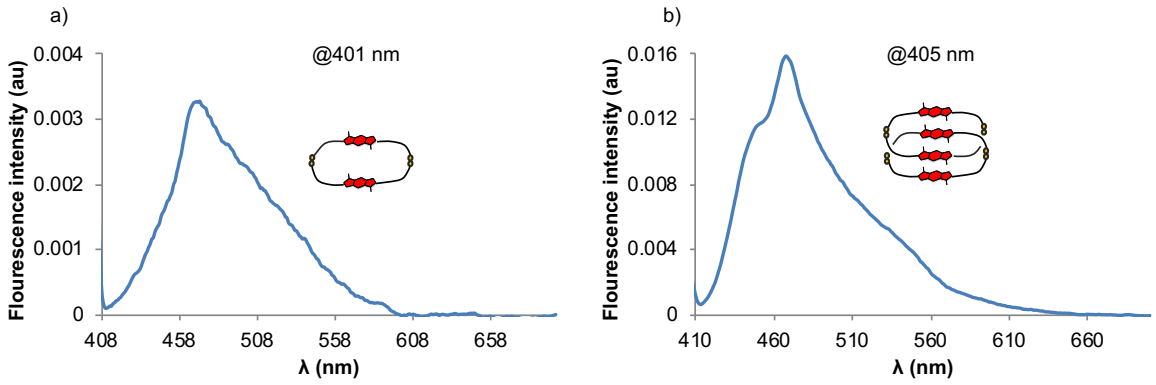


Figure 3.23. Fluorescence emission spectra in water of **Y** and ***DDDD* [2]catenane *RRRR*** (the excitation wavelength are provided on each spectrum).

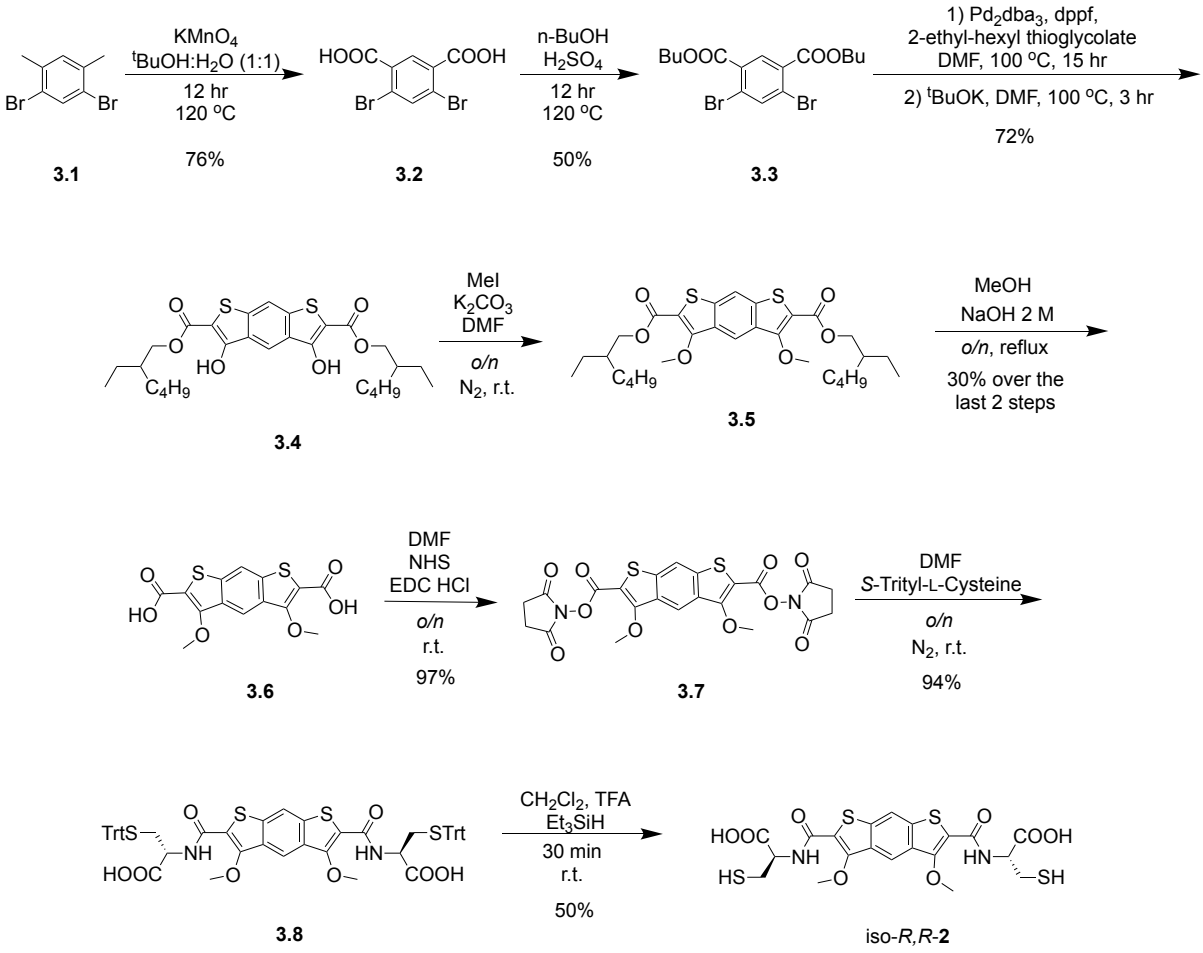
Lastly, computational studies on both homodimer (**Y**) and all-donor catenane (***RRRR***) provided further information. As expected, the heat of formation for the catenane is more than double than for the heterodimer, due to the interactions favouring its formation: π-π stacking and hydrophobicity of the aromatic cores (Table 3.1).

Table 3.1. The energy of formation of **Y** and ***DDDD* [2]catenane *RRRR***.

Compound	Energy (kJ/mol)
Y	-4215.87
DDDD* [2]catenane *RRRR	-8439.14

The second part of the chapter focuses on the iso-BDT derivative (an isomer of BDT, but with different positions for the sulphur atom. The synthesis was performed by a visiting student, Edwige Audibert (Summer student in Pantoș' group), under my direct guidance. The analysis of the building block / DCLs (LC-MS and UV / CD / Fluorescence) was done by myself and Edwige. Conclusions along with computational calculations (Q_{zz}) were done by myself; Anamaria Trandafir (Ph.D. student in Pantoș' group) did the ACID calculations. The synthetic protocol resembles the one for BDT,

but it involves one extra step: the oxidation of 1,5-dibromo-2,4-dimethylbenzene (Scheme 3.1).



Scheme 3.1. The synthetic scheme for iso-*R,R*-**2**.

3.4. Analysis of iso-*R,R*-2 library

DCLs of iso-*R,R*-**2** were set up by dissolving the building block in water at basic pH (8) followed by oxidation under atmospheric conditions. The library was stirred for about three days before analysis. The HPLC and LC-MS analyses showed the formation of two dimers and two tetramers (one dimer and one tetramer form clusters with CH_3CN – a common problem in MS – based on their masses and different retention times). Addition of salt (1 M concentration), changes significantly the library's distribution. The major peak elutes only with 100% acetonitrile; this suggests the formation of either a polymer or an oligomer because the dimer / tetramer are eluting at lower retention times than this (Figures 3.24 – 3.28). To prove our supposition, we

have monitored the library with and without salt by absorbance spectroscopy. Samples at the same concentration have different spectral profile in the presence or absence of salt. The DCL without salt displays sharp peaks, whereas the one with salt shows broad peaks. This indicates that, indeed, iso-*R,R*-**2** polymerises when the hydrophobic effect is increased (Figure 3.29).

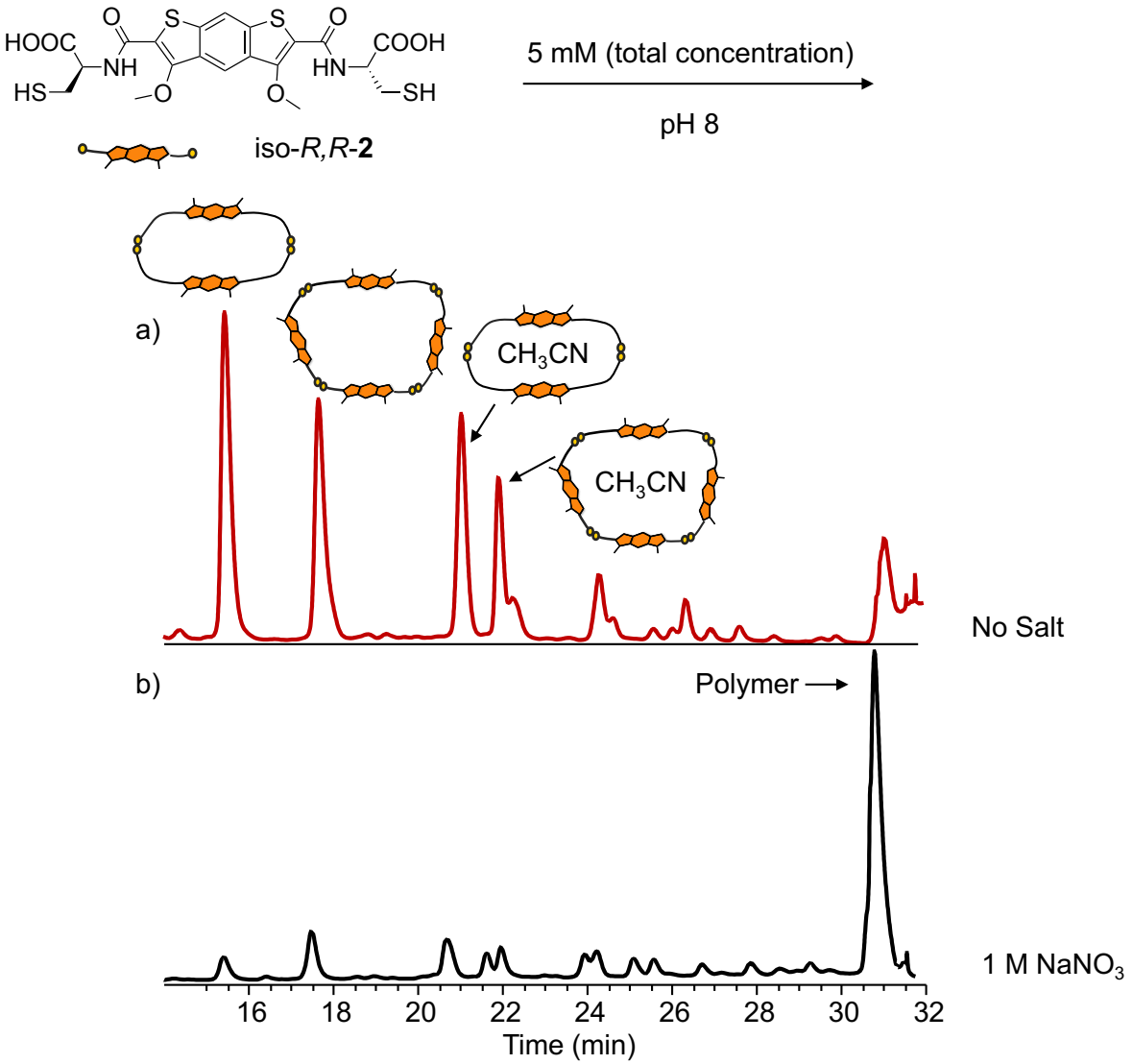


Figure 3.24. Reverse-phase HPLC analysis of iso-*R,R*-**2** (5 mM concentration) library a) without salt and b) in the presence of 1 M $NaNO_3$. Absorbance was recorded at 326 nm.

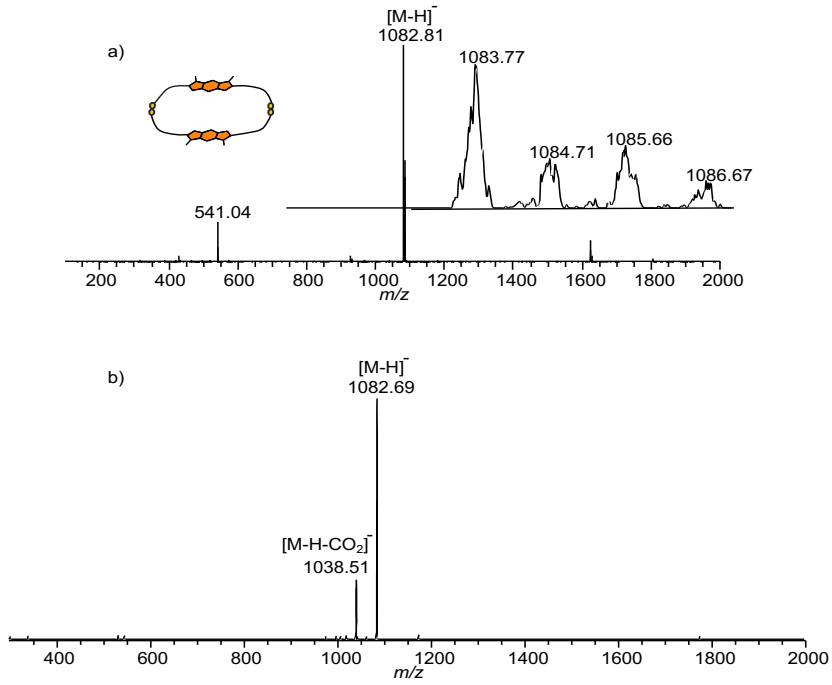


Figure 3.25. a) MS (-*ve*) of iso-*R,R*-**2** dimer; the zoom of the molecular ion is shown as an inset and b) MS/MS (-*ve*) of iso-*R,R*-**2** dimer.

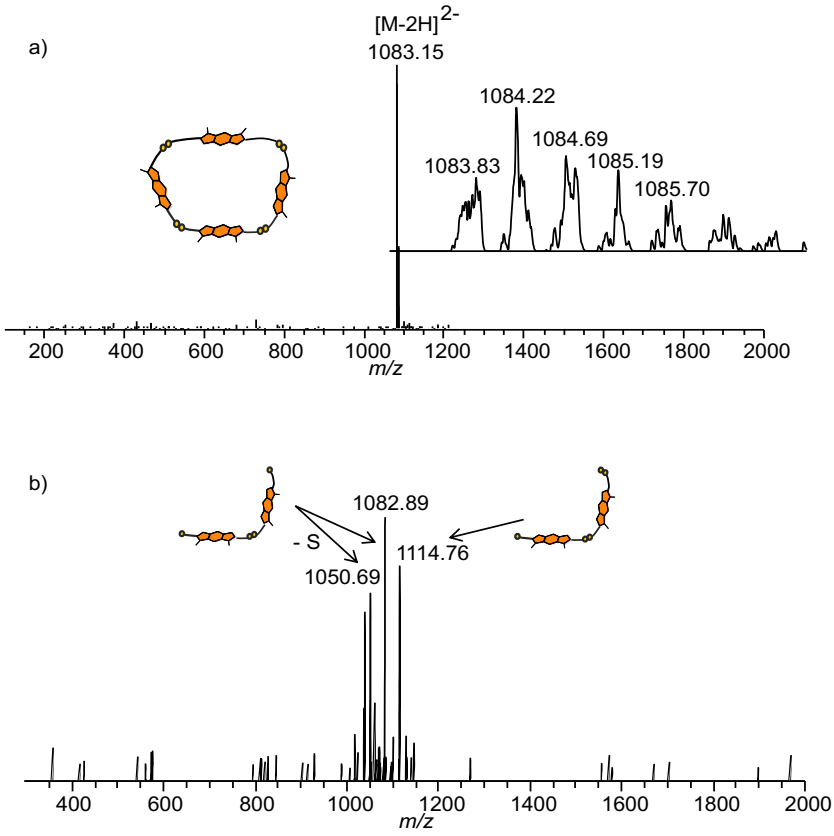


Figure 3.26. a) MS (-*ve*) of iso-*R,R*-**2** tetramer; zoom of the molecular ion is shown as an inset and b) MS/MS (-*ve*) of iso-*R,R*-**2** tetramer.

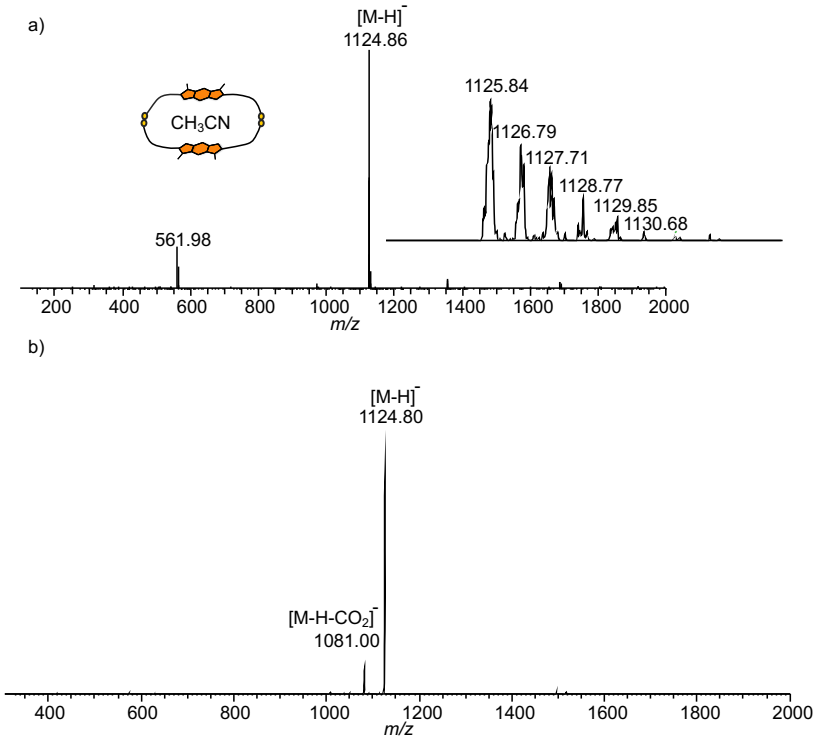


Figure 3.27. a) MS (-*ve*) of iso-*R,R*-**2** dimer cluster with a molecule of CH_3CN; the zoom of the molecular ion is shown as an inset and b) MS/MS (-*ve*) of iso-*R,R*-**2** dimer cluster with a molecule of CH_3CN.

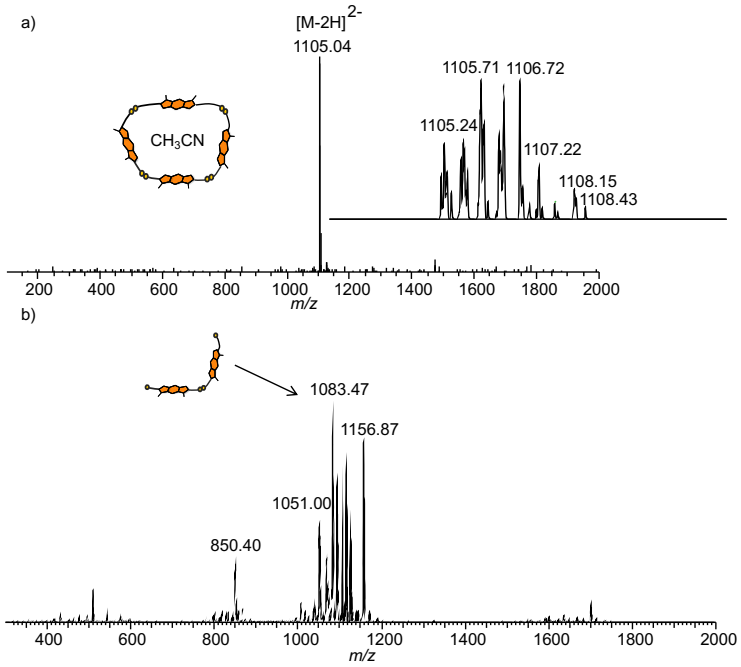


Figure 3.28. a) MS (-*ve*) of iso-*R,R*-**2** tetramer cluster with a molecule of CH_3CN; the zoom of the molecular ion is shown as an inset and b) MS/MS (-*ve*) of iso-*R,R*-**2** tetramer cluster with a molecule of CH_3CN.

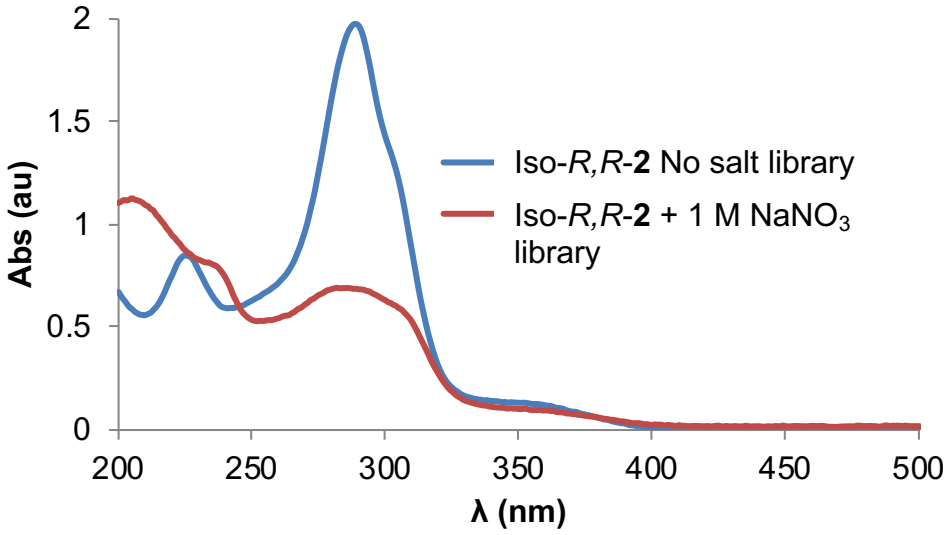


Figure 3.29. The absorbance spectra in water of iso-*R,R*-**2** libraries without salt and in the presence of 1 M $NaNO_3$.

In order to explain the behaviour of this building block, a comparison of the properties of the iso-*R,R*-**2** and *R,R*-**2** precursors is required. The absorption profile of *R,R*-**2** precursor shows a bathochromic shift compared to that of iso-*R,R*-**2** (λ_{max} 393 nm for BDT *vs.* 352 nm for iso-BDT), suggesting a higher band gap for iso-BDT-based compounds compared to BDT (Figures 3.30 and 3.31). The BDT derivatives are more fluorescent than their iso-BDT counterparts (Figures 3.32 and 3.33) – QY for **2.9** is 0.075, whereas for **3.8** it is 0.0079. The chiroptical properties of these molecules differ as well: **3.8** has a larger Cotton effect than **2.9** (Figure 3.34).

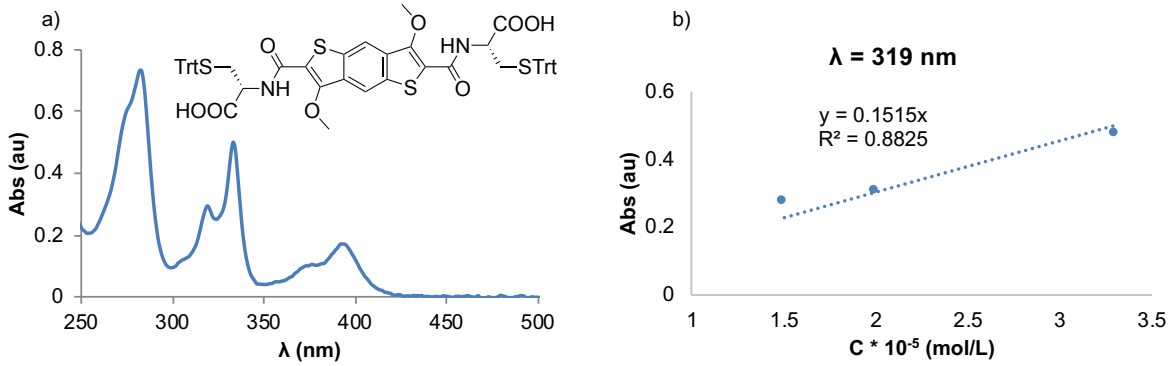


Figure 3.30. a) The structure of **2.9** and its absorbance spectrum in $CHCl_3$ and b) Lambert-Beer plot for compound **2.9** at 319 nm.

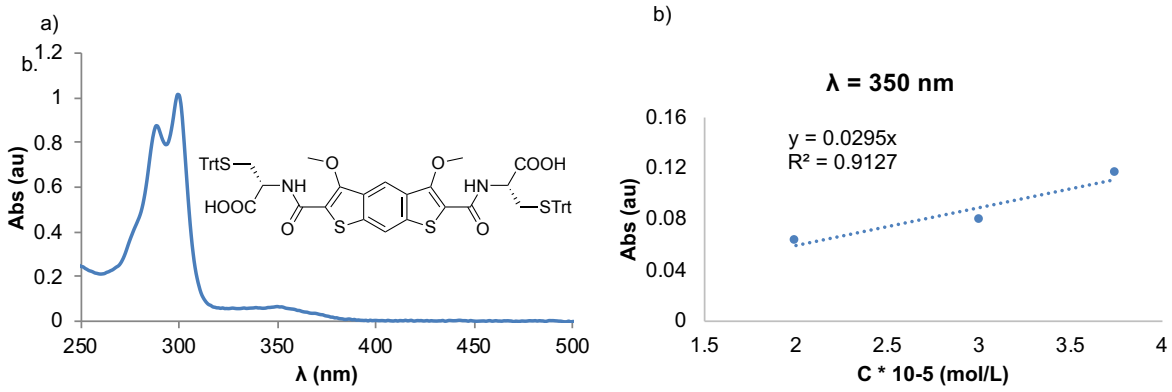


Figure 3.31. a) The structure of **3.8** and its absorbance spectrum in $CHCl_3$ and b) Lambert-Beer plot for compound **3.8** at 350 nm.

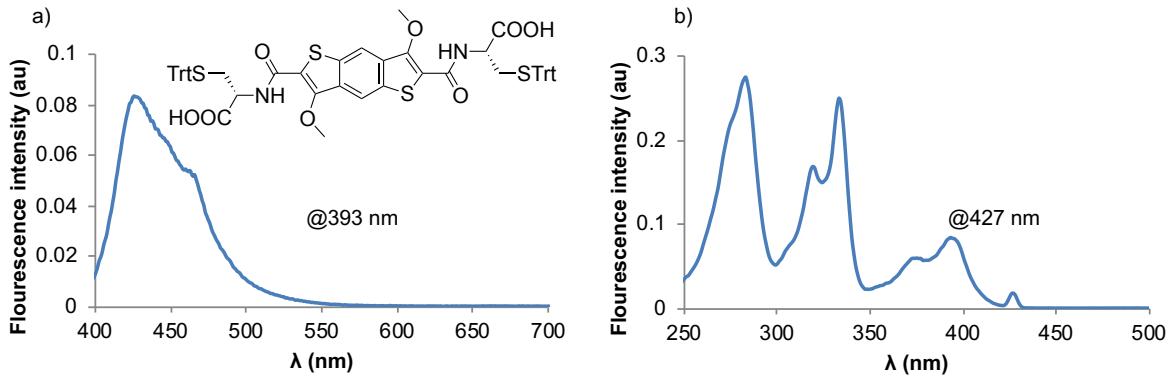


Figure 3.32. a) The emission spectrum in $CHCl_3$ of compound **2.9** (excitation wavelength 393 nm) and b) the excitation spectrum of compound **2.9** (emission wavelength 427 nm).

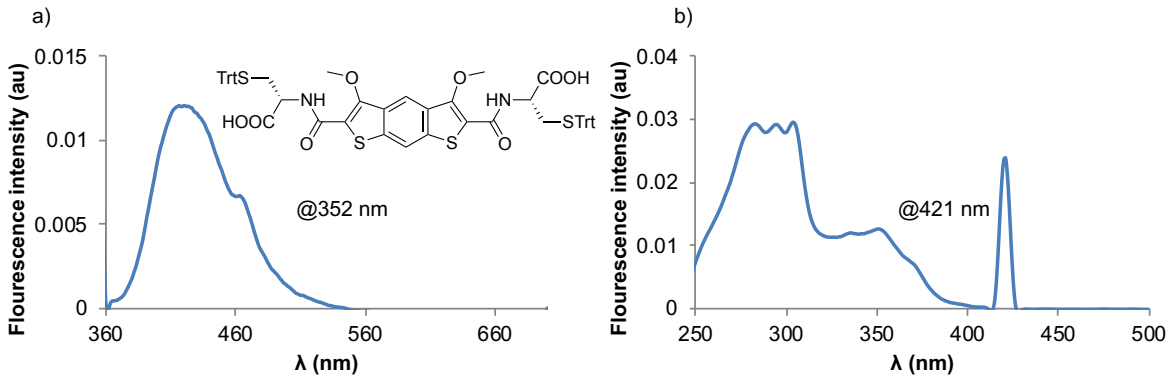


Figure 3.33. a) The emission spectrum in $CHCl_3$ of compound **3.8** (excitation wavelength 352 nm) and b) the excitation spectrum of compound **3.8** (emission wavelength 421 nm).

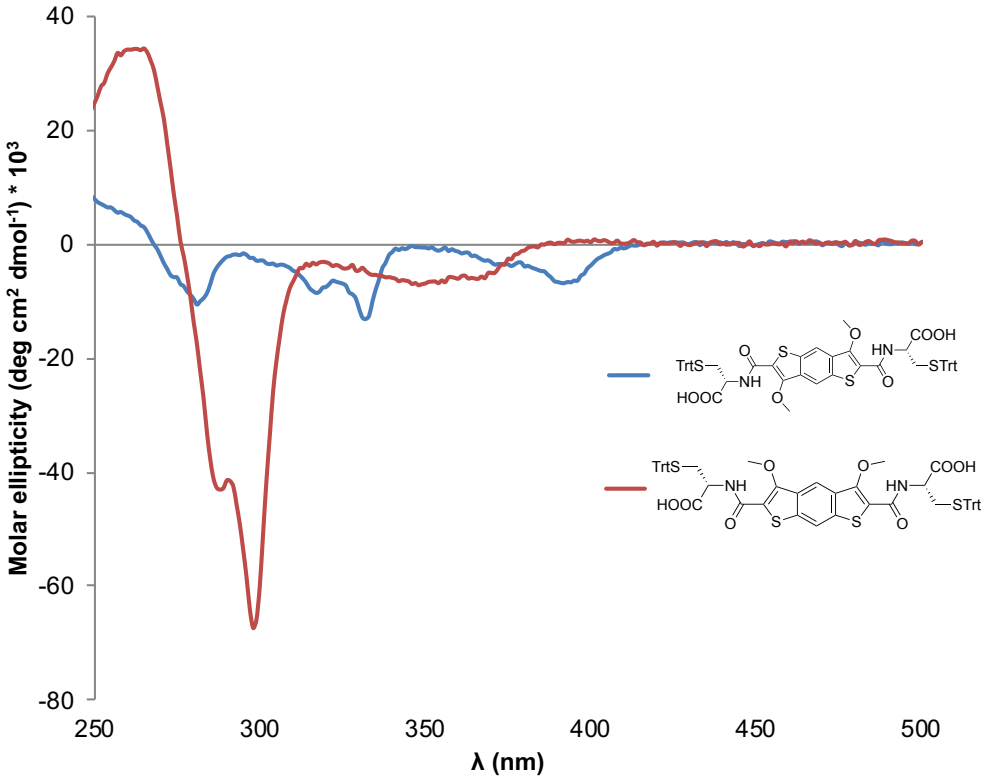


Figure 3.34. The circular dichroism spectra (in molar ellipticity) of compounds **2.9** and **3.8** in $CHCl_3$.

Lastly, we have performed computational studies on iso-BDT. The Q_{zz} for the iso-BDT core is -22 B, which is comparable with the quadrupole moment for BDT (Q_{zz} = -23 B). The electron delocalisation in a molecule is usually analysed by AICD calculations, which can show the flux of electrons in π-rich molecules. In all cases (Figure 3.35), the compounds are fully conjugated and aromatic – shown by diatropic ring current (clockwise). The bottom part of Figure 3.35 represents BDT and iso-BDT cores with amide attachment to better resemble the real system. On this basis, we can conclude that the distinct behaviour of iso-*R,R*-**2** is due to its geometry, which is slightly different from *R,R*-**2**. This very different outcomes for these isomers show the powerful nature of DCC, its selective approach, and how narrow the barrier between interlocked and non-interlocked chemistry is.

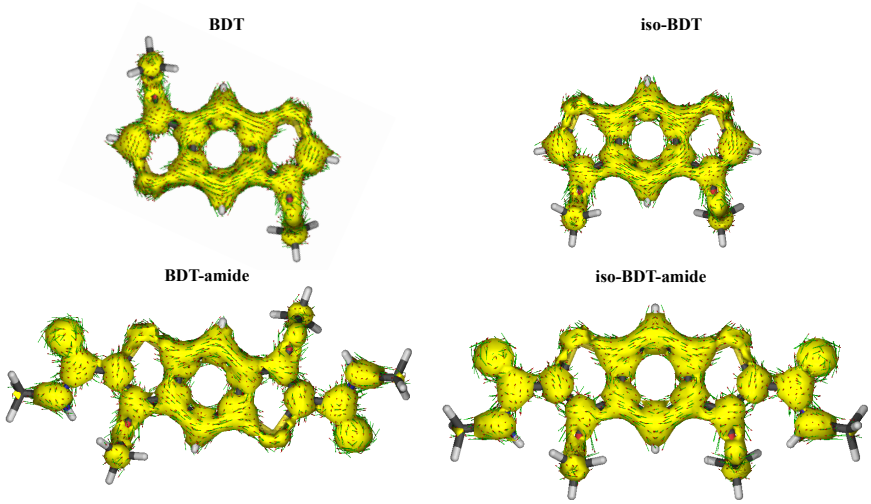


Figure 3.35. The AICD representation of BDT and iso-BDT cores (top) and BDT-amide and iso-BDT amide (bottom) (isovalue = 0.055).

3.5. Conclusion

In conclusion, this chapter reports the synthesis and characterisation of the first ever aromatic all-donor [2]catenane based on a benzodithiophene core using the DCC approach. The key step for its formation was the extended aromatic core (a quasi-pentacyclic one) along with the increased hydrophobic effect, yielding the catenane in 15% yield. The enantiomeric [2]catenane formation was formed as well as the heterochiral all-donor [2]catenane. The catenane also maintains its chiroptical properties, and because of the inherent fluorescence of the building block, the all-donor [2]catenanes are the first reported emissive catenanes synthesised by disulphide-DCC. Geometry is another key feature in the catenane formation as observed in the case of BDT *vs.* iso-BDT. BDT features a “Z”-type arrangement (*i.e.* zig-zag pattern). On the other hand, a “U”-type arrangement is observed in the case of iso-BDT.

3.6. References

1 E. Wasserman, *J. Am. Chem. Soc.*, 1960, **82**, 4433–4434.

2 E. Wasserman, *Sci. Am.*, 1962, 13.

3 G. Gil-Ramírez, D. A. Leigh and A. J. Stephens, *Angew. Chem. Int. Ed.*, 2015, **54**, 6110–6150.

4 H. Y. Au-Yeung, G. D. Pantoş and J. K. M. Sanders, *J. Am. Chem. Soc.*, 2009, **131**, 16030–16032.

5 H. Y. Au-Yeung, F. B. L. Cougnon, S. Otto, G. D. Pantoş and J. K. M. Sanders, *Chem. Sci.*, 2010, **1**, 567–574.

6 F. B. L. Cougnon, N. Ponnuswamy, N. A. Jenkins, G. D. Pantoş and J. K. M. Sanders, *J. Am. Chem. Soc.*, 2012, **134**, 19129–19135.

7 J. C. Barnes, A. C. Fahrenbach, D. Cao, S. M. Dyar, M. Frasconi, M. A. Giesener, D. Benitez, E. Tkatchouk, O. Chernyashevskyy, W. H. Shin, H. Li, S. Sampath, C. L. Stern, A. A. Sarjeant, K. J. Hartlieb, Z. Liu, R. Carmieli, Y. Y. Botros, J. W. Choi, A. M. Z. Slawin, J. B. Ketterson, M. R. Wasielewski, W. A. Goddard and J. F. Stoddart, *Science*, 2013, **339**, 429–433.

8 W. Wang, L. Wang, B. J. Palmer, G. J. Exarhos and A. D. Q. Li, *J. Am. Chem. Soc.*, 2006, **128**, 11150–11159.

9 I. R. Fernando, M. Frasconi, Y. Wu, W.-G. Liu, M. R. Wasielewski, W. A. Goddard and J. F. Stoddart, *J. Am. Chem. Soc.*, 2016, **138**, 10214–10225.

Chapter 4. NDI-BDT-NDI trimer: accessing higher-order interlocked structures?

The chemistry using DCC has led to unexpected results starting from trivial building blocks. However, if the building block has an aromatic unit with two thiol appendages, there is a considerable probability the library would contain either macrocycles or [2]catenanes.[1–6] Two aromatic units linked together by an aliphatic moiety with peripheral thiol groups would give a more complex library distribution. Previous studies reported in the literature have showed that such building blocks form [3]catenanes and also some higher-order interlocked structures: figure-eight knot, trefoil knot and Solomon link.[7,8] Literature also reports the use of a trithiol building block to make capsules. However, the complex mixture of these capsules and their similarities do not allow for separation and further characterisation.[9,10] A dithiol containing three identical aromatic units (NDI – π-acceptor molecule) was reported by the Sanders' group.[11] This building block remarkably assembles into a trefoil knot almost quantitatively (Figure 4.1 top). Similar building blocks have been previously synthesised at the University of Cambridge. In that case, the unit in the middle was a DN unit (both 1,5 and 2,6 isomers) instead of NDI. However, this trimer formed only macrocycles and unconventional [2]catenanes in low yields.[12] We have been inspired by this idea and have designed and synthesised a new trimer composed of two π-acceptor units at the periphery and one π-donor moiety in the centre, as shown in Figure 4.1 bottom.

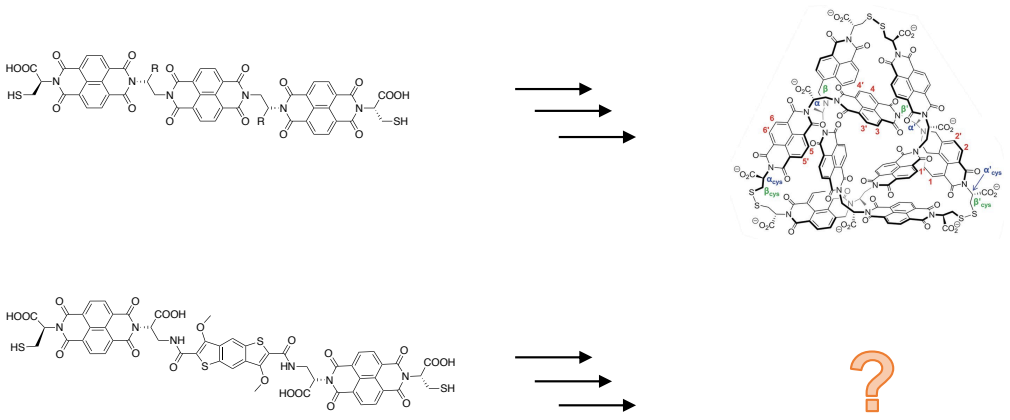
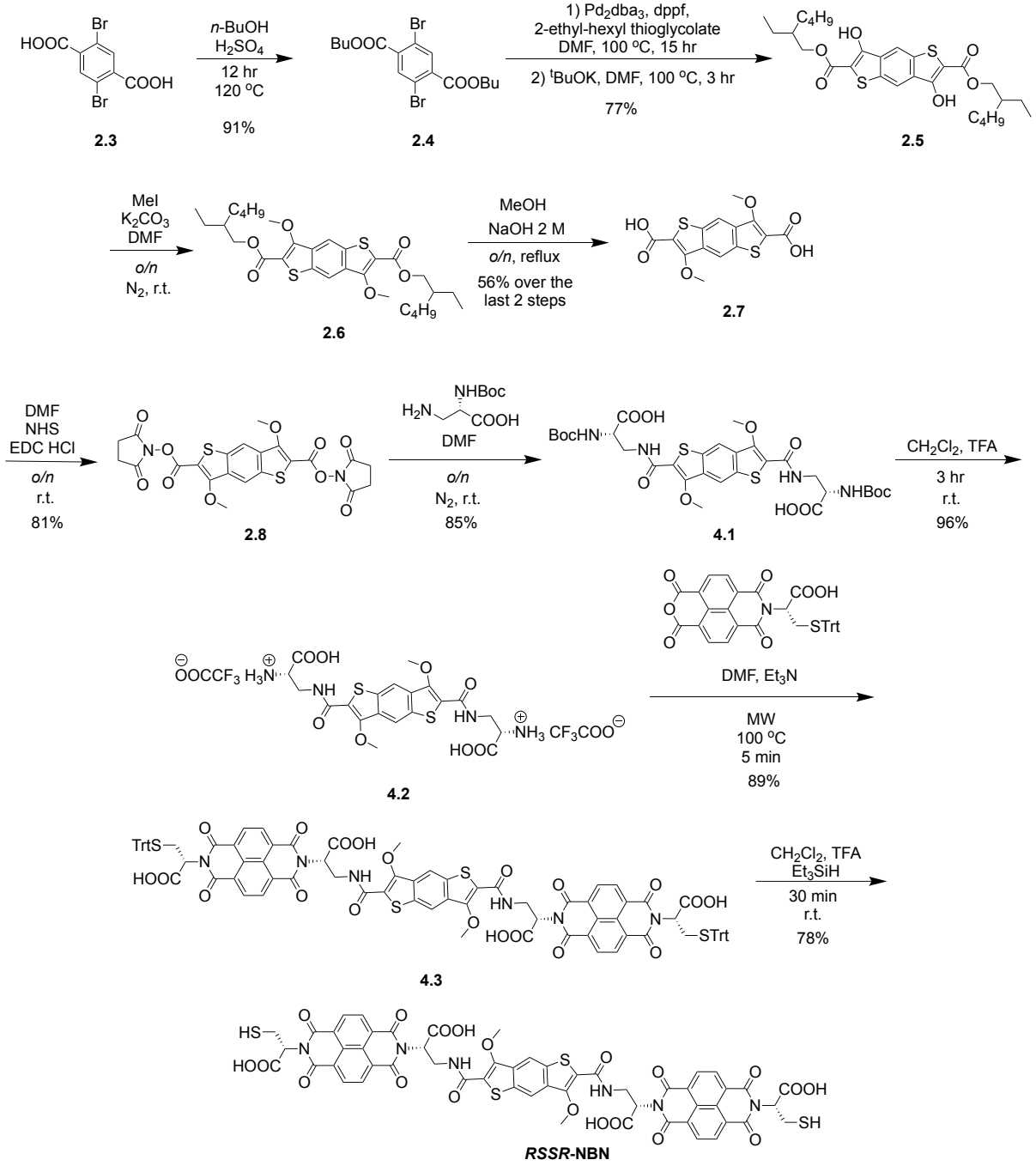


Figure 4.1. The structure of the previously synthesised trefoil knot (top), and the trimer synthesised in this work (bottom).

The synthetic scheme is outlined in Scheme 4.1. The synthesis has followed a pathway similar to the one for *R,R*-**2** until the NHS activation of BDT (**2.8**). The next step was the reaction of **2.8** with β-amino-Boc-L-alanine, which upon Boc deprotection afforded the free amine derivative **4.2** as a trifluoroacetate quaternary amine. The reaction of **4.2** with NMI-Trt-*R*-Cys gave the trimer with trityl protected appendages. The last step was the deprotection, yielding the red-coloured ***RSSR*-NBN**.



Scheme 4.1. The synthetic scheme for ***RSSR*-NBN**.

4.1. Analysis of *RSSR*-NBN library

DCLs (with and without additives) were prepared by dissolving ***RSSR*-NBN** in NaOH 10 mM, and the pH was adjusted to 8 using 100 mM NaOH / 100 mM HCl. The distribution of the library with no salt or additives showed a monomer (12%), two dimers (61%) and a tetramer (27%). Initial studies (MS, zoom scan and MS/MS) indicated that one dimer is a macrocycle. The fast retention time for the peak corresponding to the tetramer (the second species to elute) along with the MS/MS profile has suggested an interlocked species – labelled as "knot"; the same is true for the other dimer (the catenated structure – it will be analysed in depth later in this chapter). Interestingly, when 1 M $NaNO_3$ was added to increase the hydrophobic effect, the system was composed of the tetramer as the major species along with the monomer and the macrocyclic dimer. Figures 4.2 – 4.6 show the detailed analysis of these species (the zoom scan for the tetramer is high resolution). In contrast, when the library was made in the presence of 1 M $TBANO_3$, all the interlocked species disappeared.

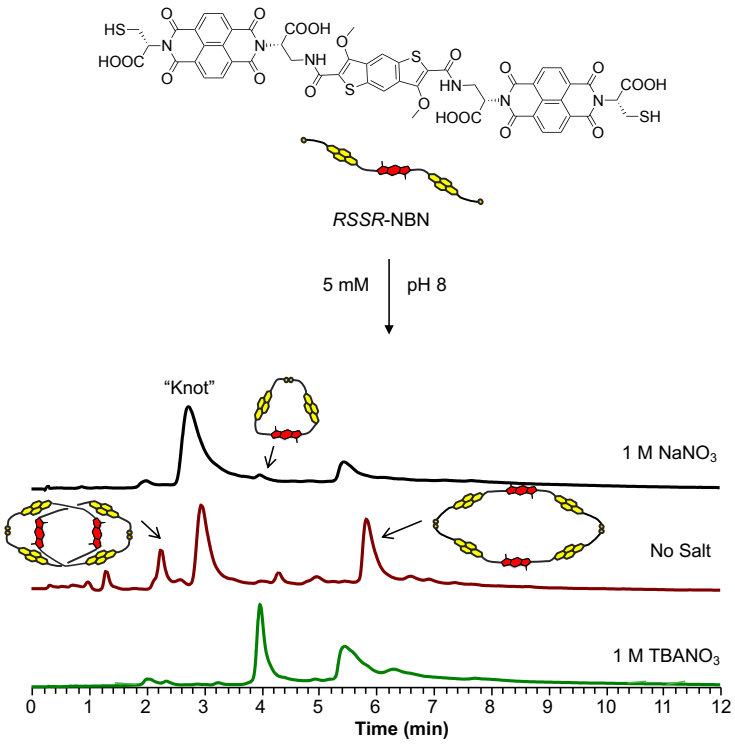


Figure 4.2. Reverse-phase HPLC analysis of ***RSSR*-NBN** (5 mM total concentration) library in the presence of 1 M $NaNO_3$, no salt, and in the presence of 1 M $TBANO_3$. Absorbances recorded at 389 nm. The unlabelled peaks did not ionise and could not be identified.

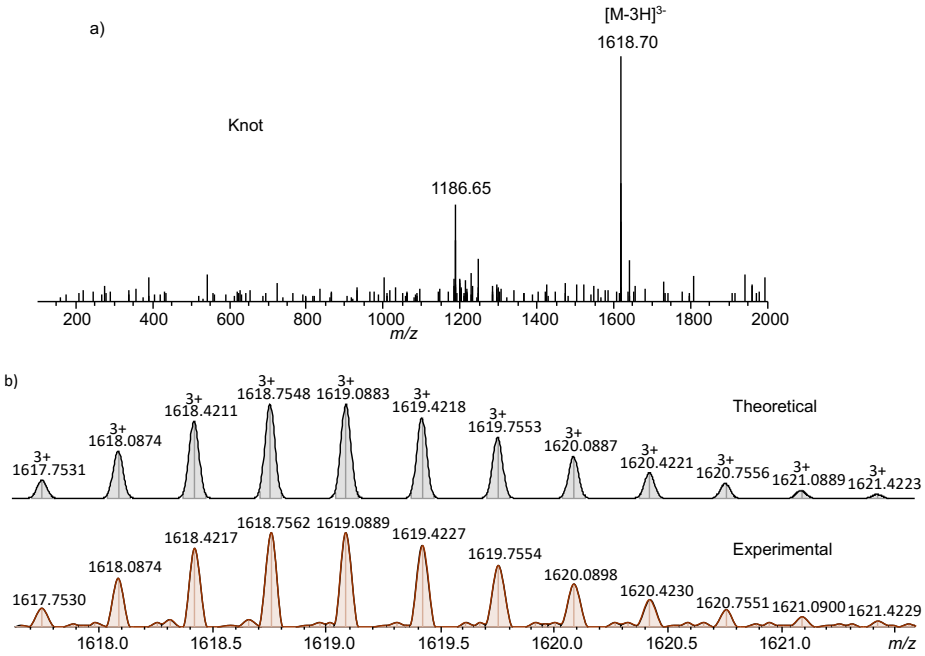


Figure 4.3. a) MS (-*ve*) of ***RSSR*-NBN knot** and b) zoom scan of the molecular ion.

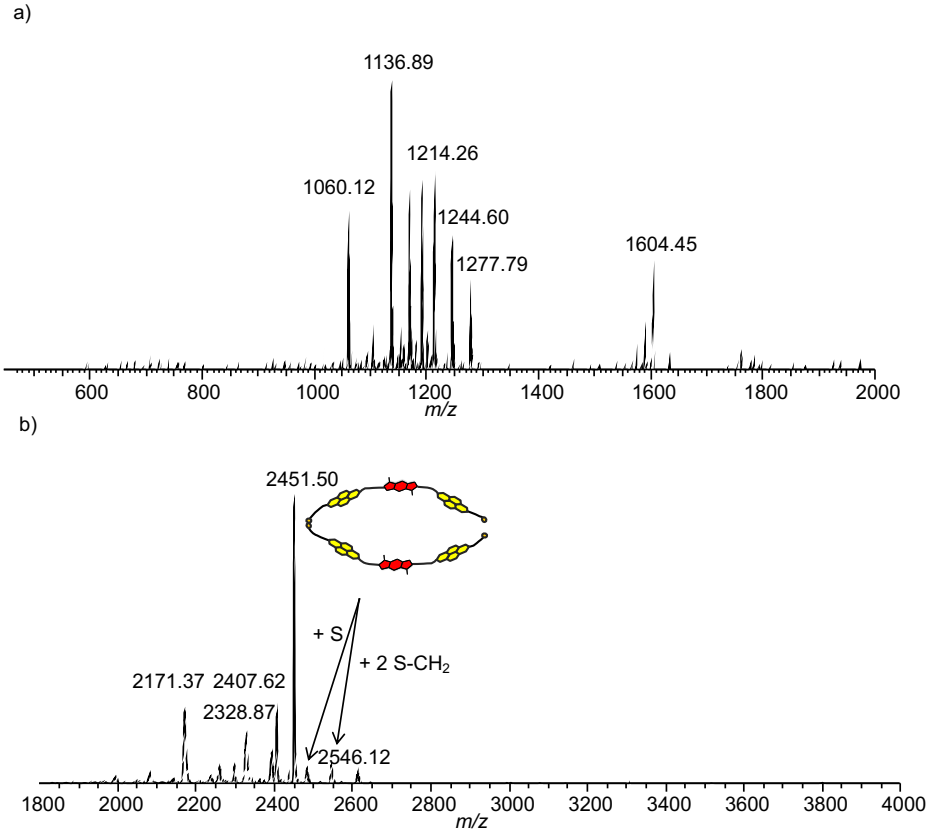


Figure 4.4. a) MS/MS (-*ve*) of ***RSSR*-NBN knot** and b) MS/MS (-*ve*) of ***RSSR*-NBN knot** between 1800 and 4000 *m/z*.

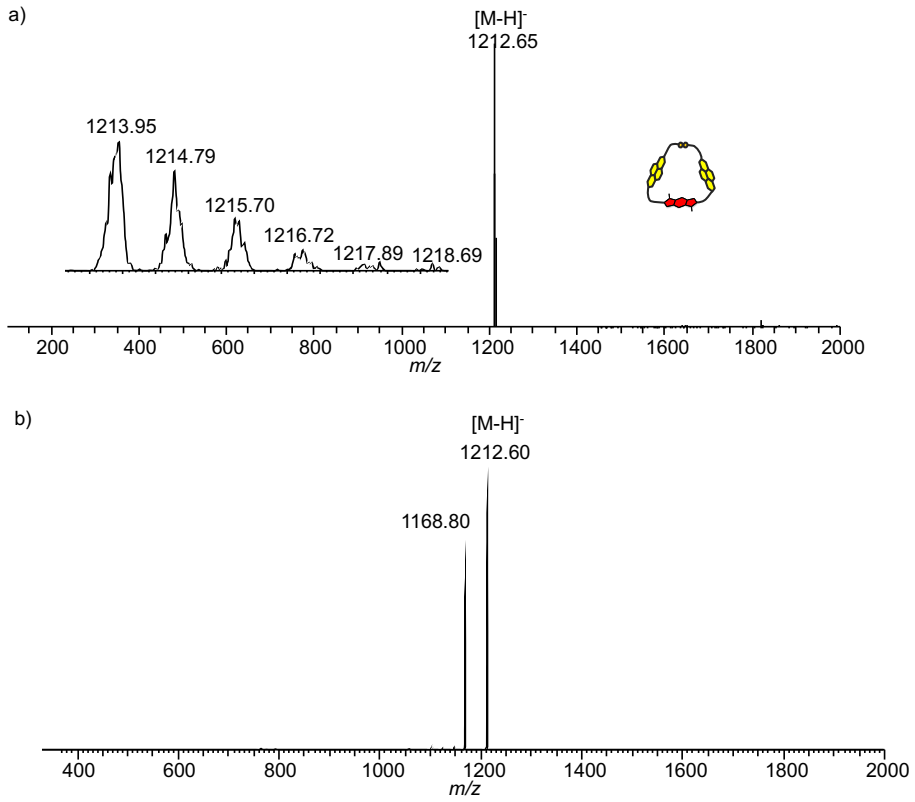


Figure 4.5. a) MS (-*ve*) of ***RSSR*-NBN monomer**; zoom of the molecular ion is shown as an inset and b) MS/MS (-*ve*) of ***RSSR*-NBN monomer**.

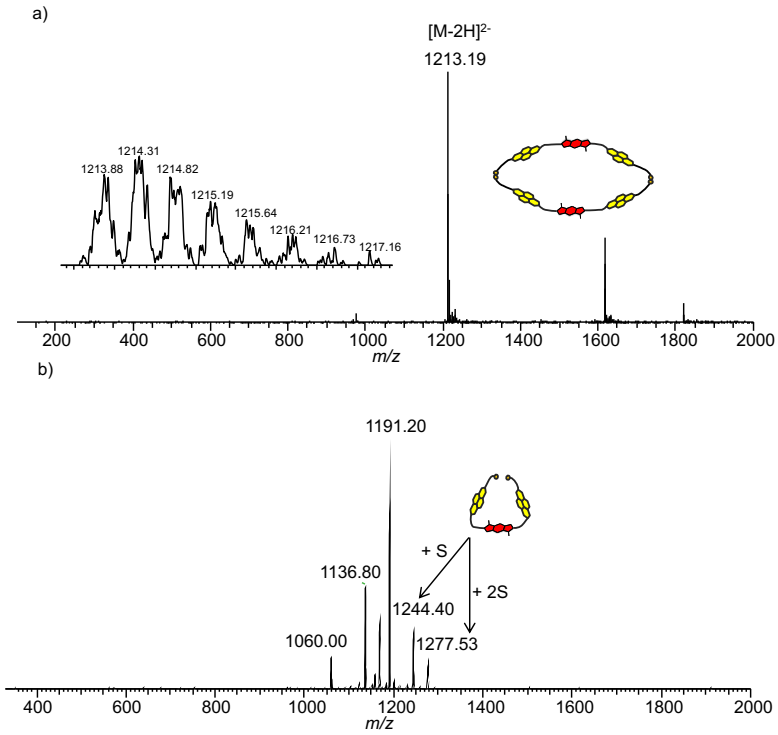


Figure 4.6. a) MS (-*ve*) of ***RSSR*-NBN dimer**; zoom of the molecular ion is shown as an inset and b) MS/MS (-*ve*) of ***RSSR*-NBN dimer**.

4.1.1. Analysis of the knot

This sparked our interest and the DCL was redone in D_2O, which was analysed using NMR techniques. The analysis of the ^1H NMR spectrum of the library shows a complex pattern, with peaks corresponding to aromatic protons ranging from 9.3 to 6.2 ppm. There are four singlets in the aromatic region identified as part of the BDT unit. This assumes that each unit of the tetramer is distinct on the NMR timescale. Figure 4.7 displays the partial ^1H NMR spectrum of this molecule, showing the aromatic region (the peaks are labelled based on the COSY spectrum; A1 – A8 and A1' – A8' correspond to NDI protons and D1 – D2 and D1' – D2' correspond to BDT protons. The 1 – 8 and 1' – 8' numbering is based on the cross peaks observed, for example, there is a cross peak between A1 and A1').

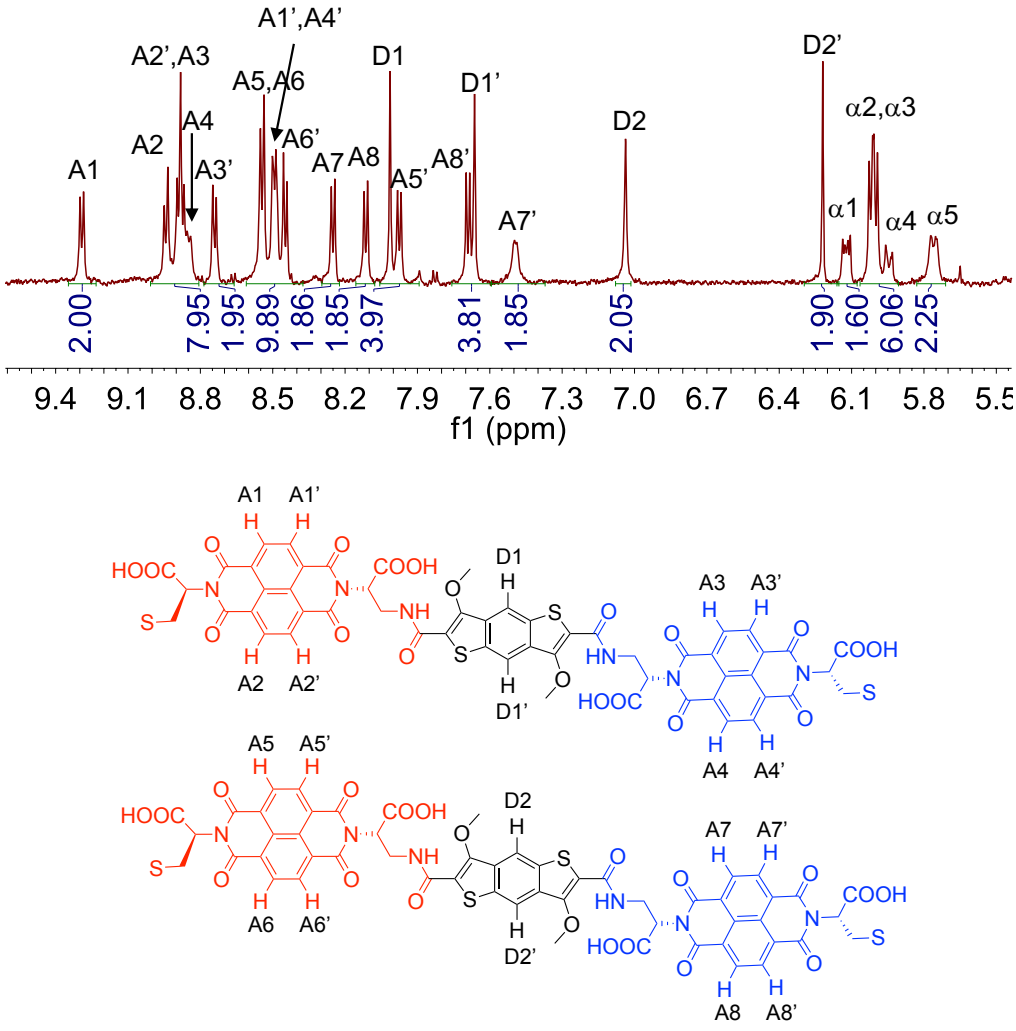


Figure 4.7. Partial ^1H NMR spectrum (500 MHz) of the library. The residual solvent (H_2O) was referenced at 4.79 ppm (top), the A1-A8 and D1-D2 possible arrangements (bottom)..

Further structural information was provided by the 2D NMR experiments: ^{1}H-^{1}H COSY, NOESY, ROESY and HSQC. Let us start by looking at the COSY analysis (Figures 4.8 and 4.9), as this shows the coupling through bonds (usually 3 – 4 bonds away, *i.e.* ^{4}J is seen in U-, W-, allylic or benzylic couplings). Based on this, there should be no cross peaks between the BDT signals. However, there are clear cross peaks between the singlets at 8.00 and 7.73 ppm, and 7.05 and 6.20 ppm (due to the very small coupling constant, the signals appear as singlets). This suggests correlation between protons that are five bonds away, which is very unusual, but not unprecedented.[13] It implies that two by two units in the tetramer are similar. For the region corresponding to the NDI protons, COSY indicates eight cross peaks, so in total sixteen signals. In other words, the NDI units are in different chemical environments, adopting either one of the two types of arrangements (due to complexity of the spectrum, the cross peaks in the aliphatic region are not labelled). Figure 4.10 shows these possible arrangements: either eight NDI cores arranged into type I (red NDI) or four NDI cores into type II (blue NDI) or a mixture of these two types. Type I represents a typical desymmetrised NDI core, with two different aromatic signals (*i.e.* H_a and H_b – Figure 4.10, red) and two α protons per NDI unit (α protons are the hydrogen atoms attached to the carbon atom in the α position with respect to the NDI core). For the proposed type II arrangement, all aromatic protons are in different chemical environments (H_a to H_d – Figure 4.10, blue) and there are two distinct α protons. The α – β connections in the aliphatic region indicate eight α protons and sixteen β protons. As there are sixteen NDI and eight α signals, the NDI cores are more likely to fall into the Type II category. Figure 4.1 summarises all possible arrangements of the aromatic protons of the acceptor-donor-acceptor unit. All these arguments indicate a structure in which each unit of molecule has eight different NDI protons, two different BDT protons, four α protons and eight β protons (first case in Figure 4.11). The HSQC analysis confirmed the above assignment, with chemical shifts of carbon atoms matching the pattern observed for both NDI and BDT units (Figure 4.12).

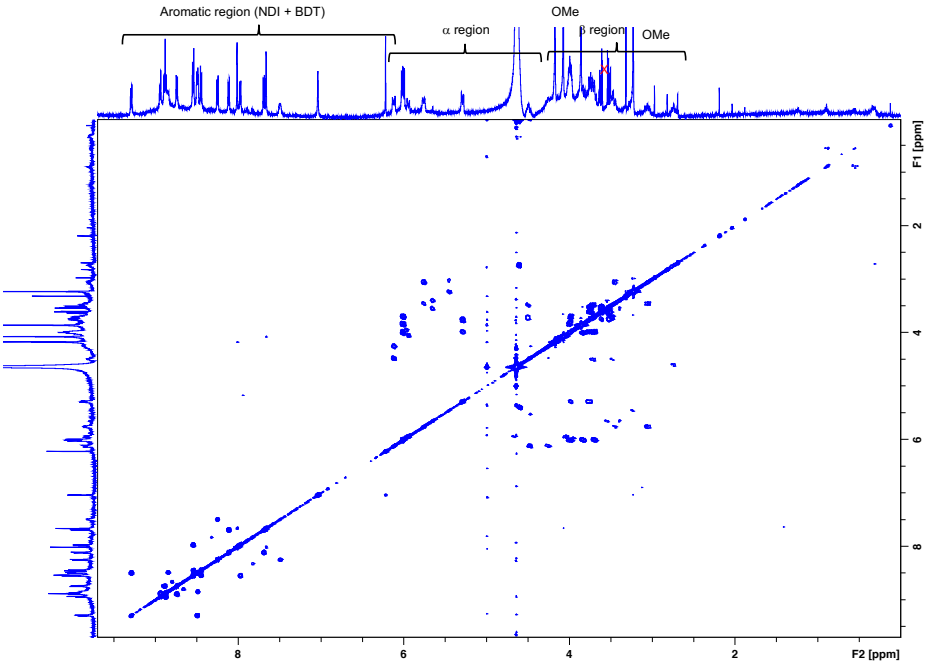


Figure 4.8. ^{1}H-^{1}H COSY spectrum (500 MHz, 298 K) of ***RSSR*-NBN knot**. The residual solvent (H_2O) was referenced at 4.79 ppm. The protons crossed with "X" are impurities.

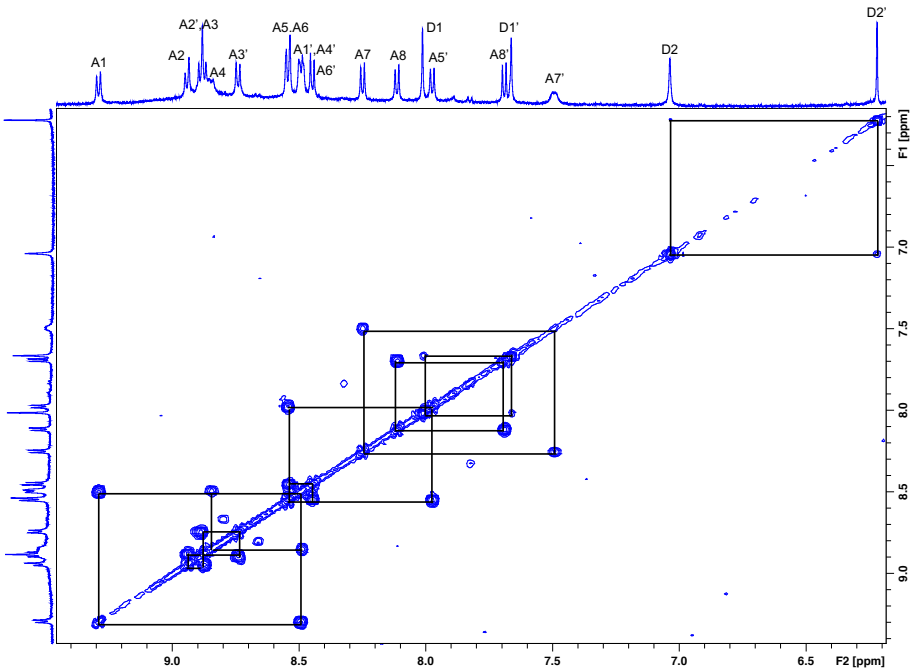


Figure 4.9. Partial ^{1}H-^{1}H COSY spectrum (500 MHz, 298 K) of ***RSSR*-NBN knot** showing the aromatic region. The residual solvent (H_2O) was referenced at 4.79 ppm (the unmarked cross peaks come from the impurities present in the library).

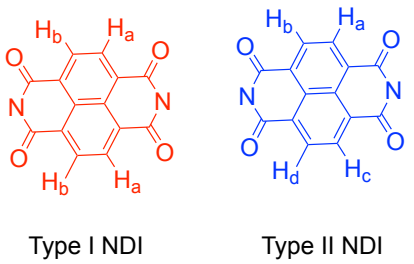


Figure 4.10. The two possible types of NDI.

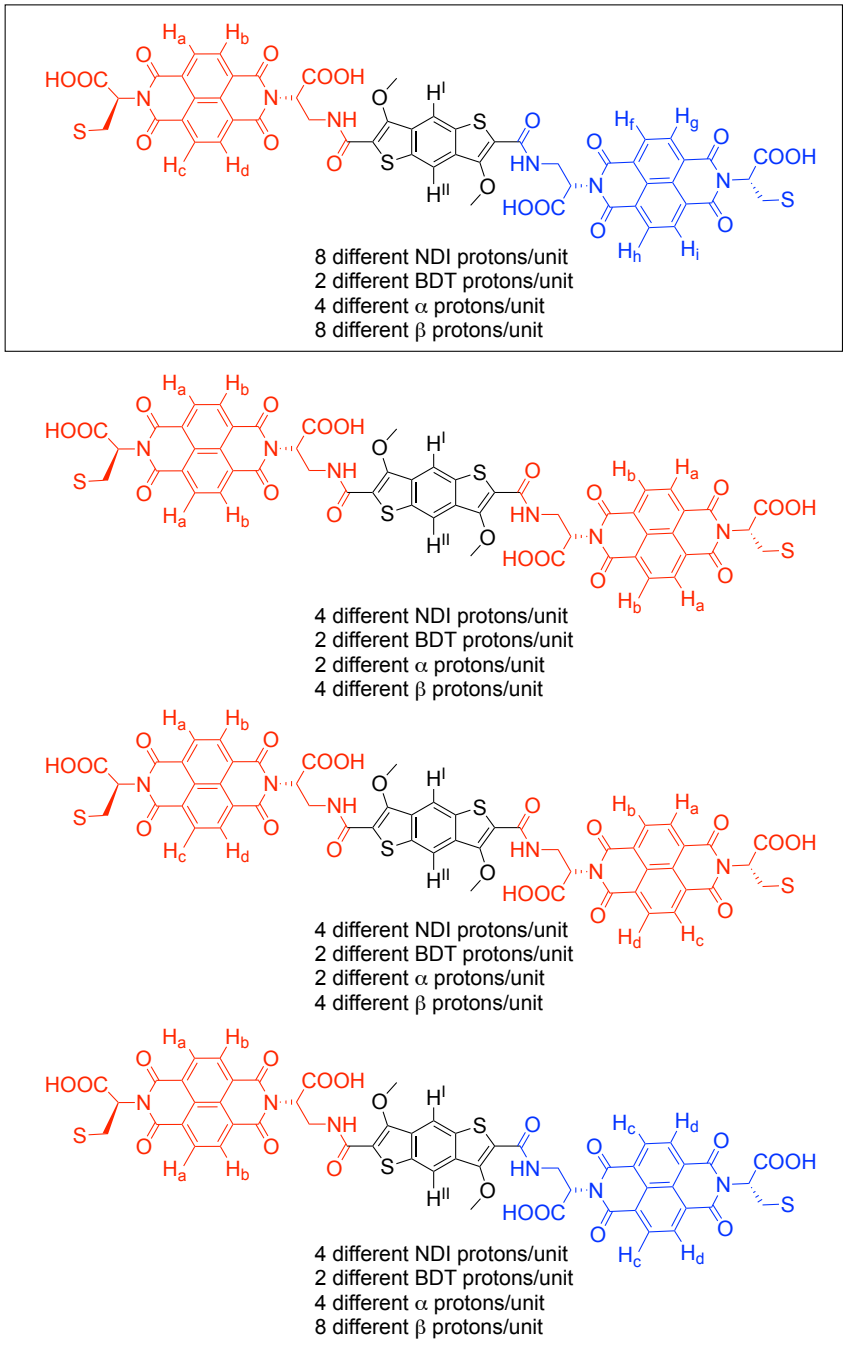


Figure 4.11. The possible arrangements of trimer aromatic protons (the one in the rectangle fits the COSY description).

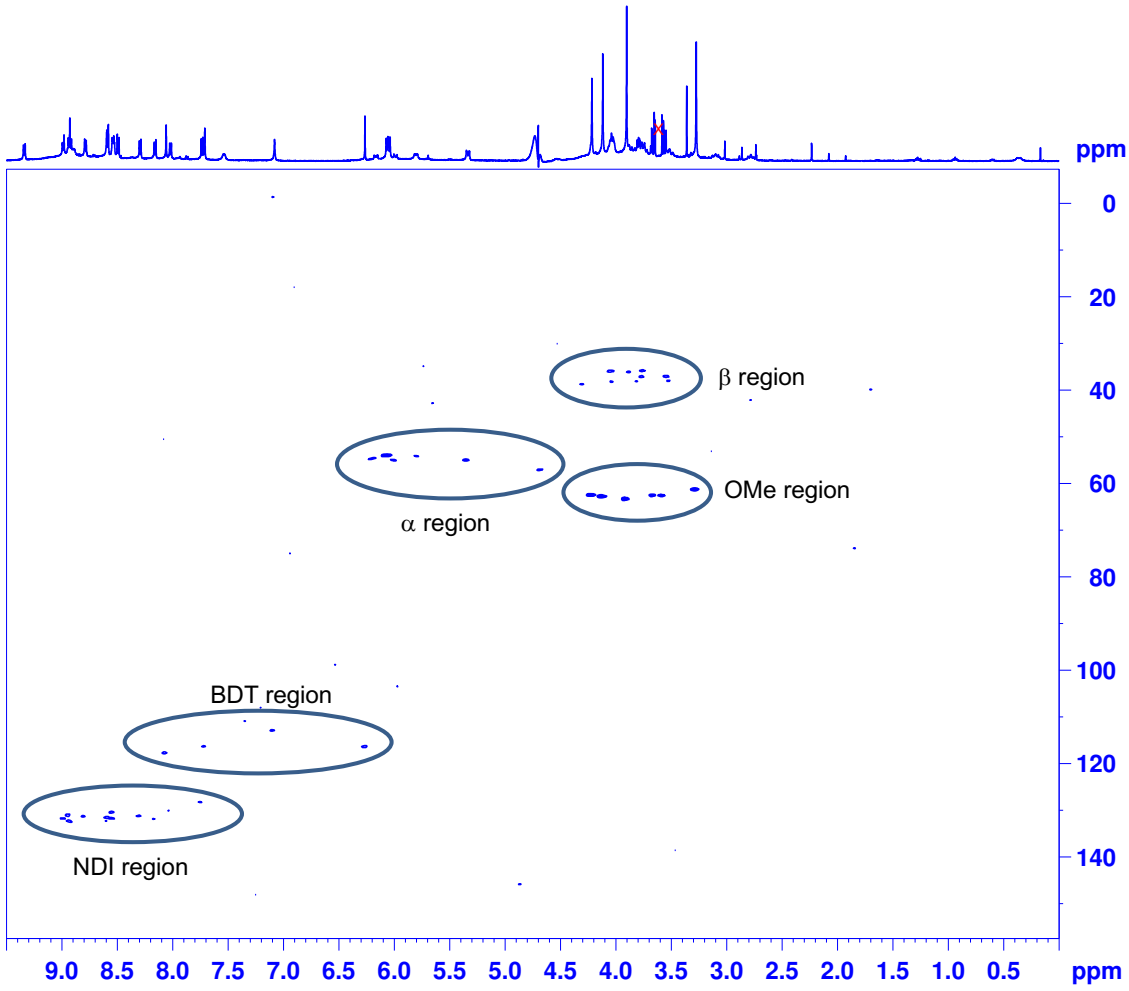


Figure 4.12. HSQC spectrum (500 MHz, 298 K) of ***RSSR*-NBN knot**. The residual solvent (H_2O) was referenced at 4.79 ppm. The protons crossed with "X" are impurities.

NOESY and ROESY experiments (Figures 4.13 – 4.16) are also useful tools for studying such a complex arrangement. The first striking observation is that NOESY and ROESY spectra do not show many cross peaks in the aromatic region. The NDI and BDT peaks show no correlation to each other. This suggests that either the cores are in a T-shape position related to each other (from a 3D point of view – see Figure 4.18) or they are not close in space. Most of the BDT peaks are close in space to β protons, and only one BDT has a cross peak with an α proton. The same is true for the NDI protons. There are also cross peaks between the methoxy groups and NDI, BDT and aliphatic moieties. VT NMR studies (Figure 4.17) have showed that some of the peaks broaden as the temperature increases, suggesting that a part of the

molecule is not as compact as the other part. Because the NMR spectrum of this molecule is very complex, the structure determination is not straightforward. NMR and MS techniques give complementary information for structure determination, allowing us the differentiate between various topologies. At this point, we have considered three main types of molecules: unknotted macrocycle, catenane and knot.

The MS and MS/MS data suggest the formation of a macrocycle, as indicated by the fragmentation pattern and the presence of fragments with *m/z* higher than half the mass of the molecule. The sharp peaks observed in the ^{1}H NMR spectra discard the unknotted macrocycle structure. Usually, unknotted macrocycles are flexible chains, adopting multiple conformations at room temperature, thus showing broad peaks in NMR (sharp peaks at room temperature are also possible for an unknotted macrocycle, but they will get very broad as the temperature increases – this is not the case of our species as most of the peaks remain sharp at high temperatures). The presence of aromatic peaks up-field are also uncharacteristic for simple macrocycles.

The next possible topology to consider in our assignment is the [2]catenane. This time, NMR analysis would be in accordance with such structure due to the sharp and shielded peaks. However, based on MS/MS data, only one type of catenane would be possible: an interlocked structure containing a monomer and a trimer. However, such a catenane would be unlikely to have a C_2 symmetry due to the flexible trimer macrocycle. Based on these arguments a catenated structure is also discarded.

The last possible structure is the knot. This type of structure is consistent with the NMR data: sharp and shielded peaks along with different environments for the NDI protons and C_2 symmetry. The MS data is also in good agreement with a knotted structure having the correct fragmentation pattern. Once the type of topology has been identified, the next step is to determine the type of knot. Based on NMR and MS/MS information, four possible knotted topologies have been proposed: the 3_1, 4_1, 5_1 and 5_2 topologies (these are Alexander-Briggs notations for knots – where the first value represents the number of crossing points and the subscript is the order of the knot in its subdivision) – Figure 4.19. Other topological arrangements are excluded because they have too many crossings for a molecule with twelve aromatic units, which lead to untenable strain in the structure. Unfortunately, 1D and 2D NMR spectra do not allow for a clear distinction between these topologies, however the topology has to have the C_2 symmetry. Many attempts to grow single crystals, including the use of a robot for crystallisation with over 400 conditions, were done, unfortunately, with no success.

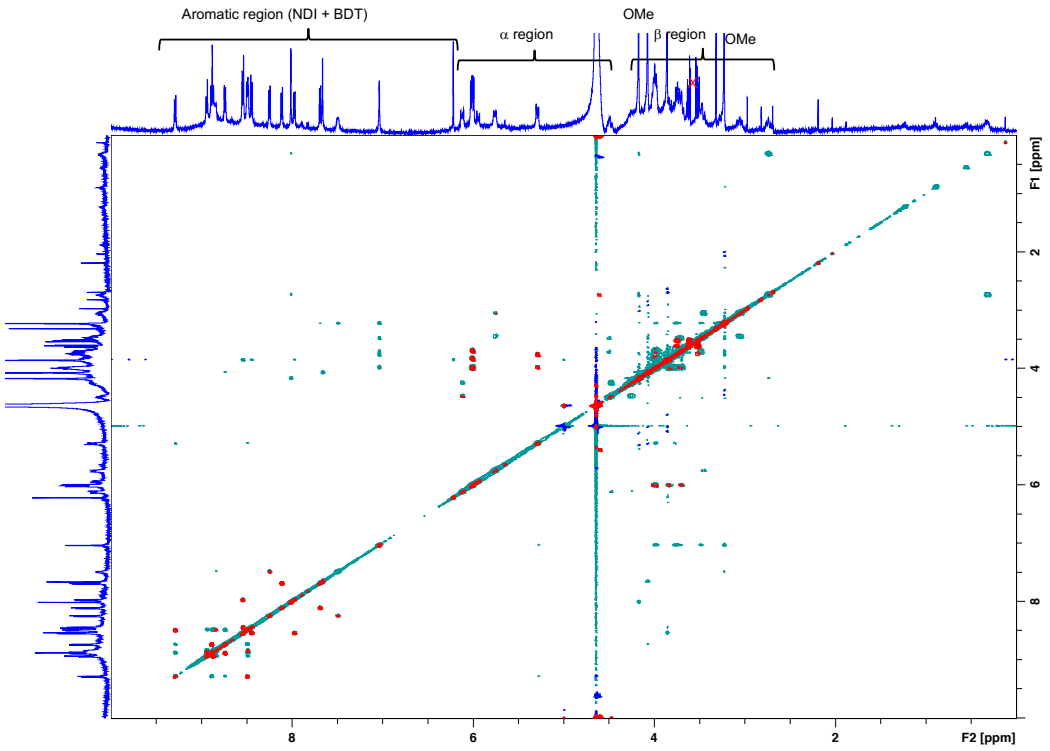


Figure 4.13. ^{1}H-^{1}H COSY (red) and 2D NOESY (green) spectra (500 MHz, 298 K) superimposed of ***RSSR*-NBN knot**. The residual solvent (H_2O) was referenced at 4.79 ppm. The protons crossed with "X" are impurities.

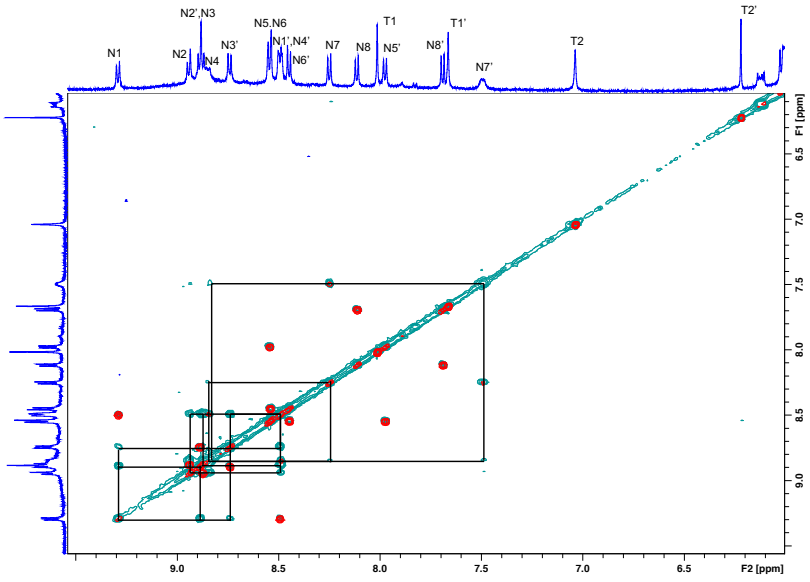


Figure 4.14. Partial ^{1}H-^{1}H COSY (red) and 2D NOESY (green) spectra (500 MHz, 298 K) superimposed of ***RSSR*-NBN knot**. The residual solvent (H_2O) was referenced at 4.79 ppm.

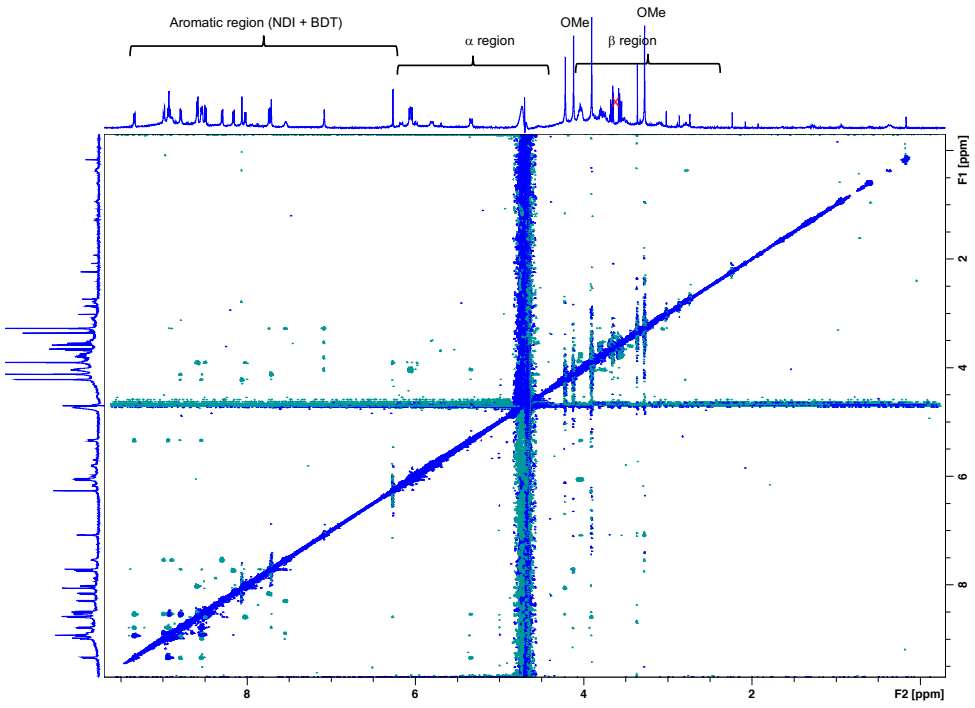


Figure 4.15. 2D ROESY spectrum (500 MHz, 298 K) of the ***RSSR*-NBN knot**. The residual solvent (H_2O) was referenced at 4.79 ppm. The protons crossed with "X" are impurities.

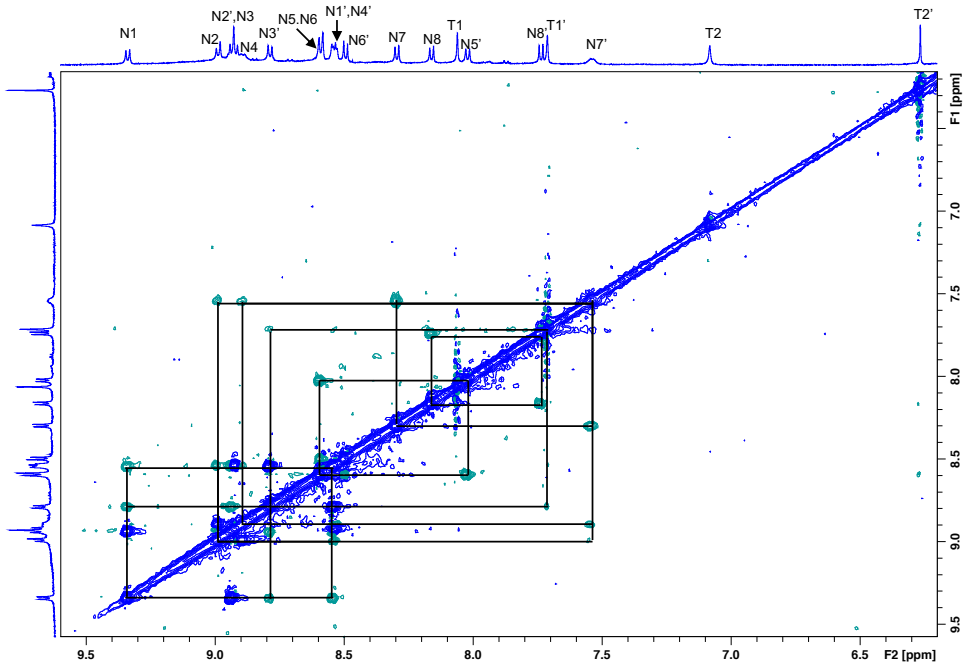


Figure 4.16. Partial 2D ROESY spectrum (green) (500 MHz, 298 K) of ***RSSR*-NBN knot**. The residual solvent (H_2O) was referenced at 4.79 ppm

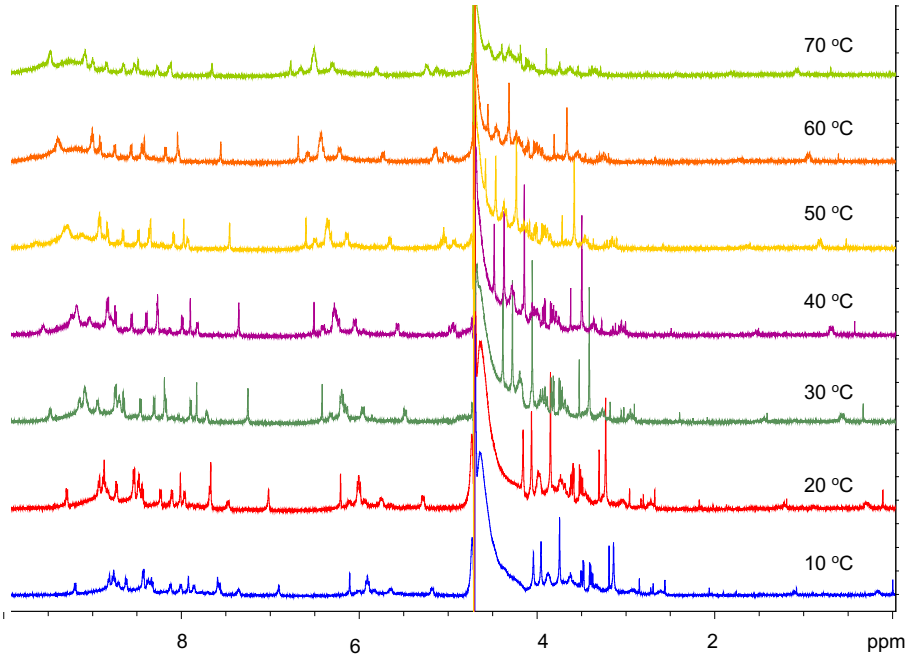


Figure 4.17. VT ^{1}H NMR (500 MHz) spectra of ***RSSR*-NBN knot** at the temperatures indicated above each spectrum. The residual solvent (H_2O) was referenced at 4.79 ppm.

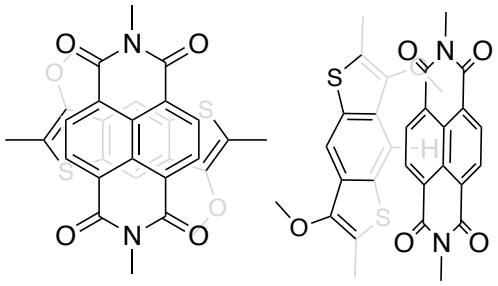


Figure 4.18. The possible arrangements of NDI and BDT cores.

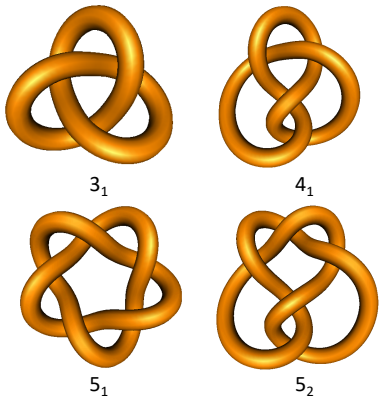


Figure 4.19. The cartoon representations of the possible knots arrangements (3_1, 4_1, 5_1 and 5_2).

The optical properties of the knot have been also investigated and have provided a useful tool in its structure determination (Figure 4.20). The UV-vis spectrum shows bands characteristic to the NDI and BDT units. The CD profile displays a relatively large Cotton effect, with bisignate bands. The small band centred around 380 nm corresponds mainly to the NDI core, while the large response is mostly given by a combination of BDT and NDI transitions (λ = 288 nm). The molecule is rigid and compact, with small changes at high temperatures, as shown by VT UV-vis and CD studies.

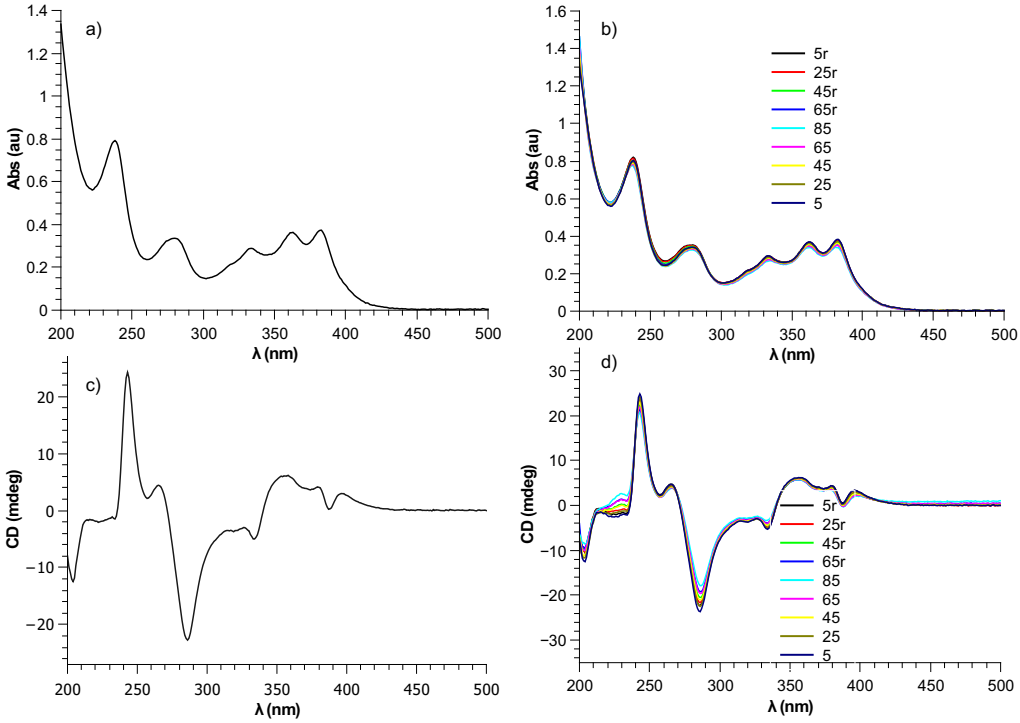
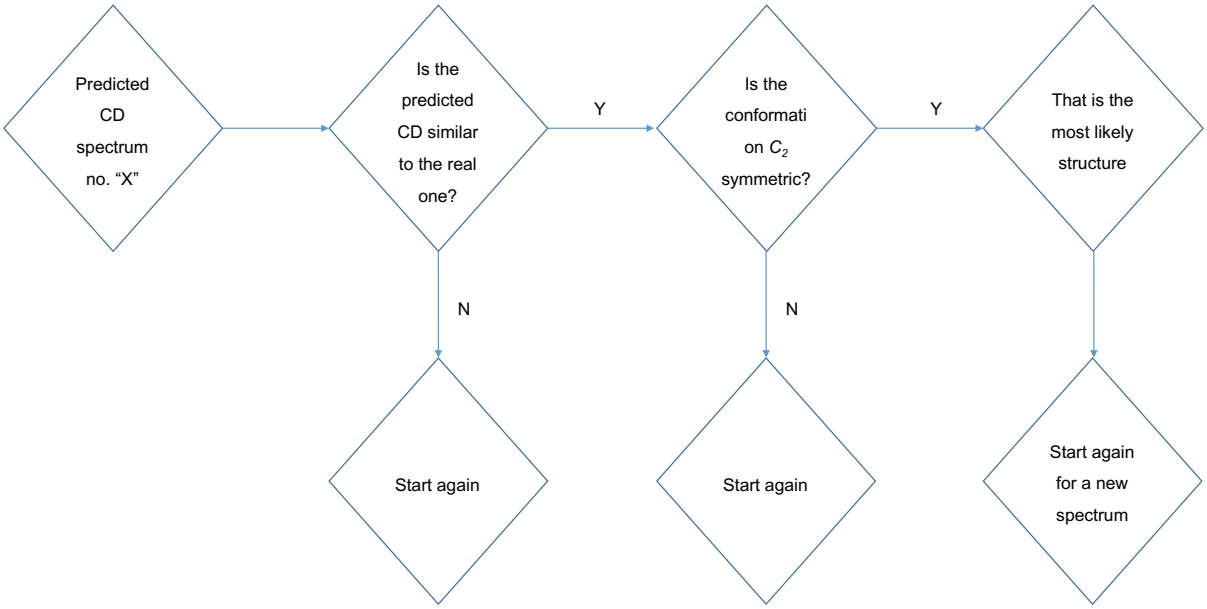


Figure 4.20. a) UV-Vis spectrum of ***RSSR*-NBN knot** at 23 °C (296 K), b) VT UV-Vis spectra of ***RSSR*-NBN knot** between 5 – 85 – 5 °C at the specified temperatures, c) CD spectrum of ***RSSR*-NBN knot** at 23 °C and d) VT CD spectra of ***RSSR*-NBN knot** between 5 – 85 – 5 °C at the specified temperatures (r – return temperature).

CD is a powerful tool for differentiating between fine structural details, and it can be used as a fingerprint for each of the possible topologies. Each possible arrangement (*i.e.* 3_1, 4_1, 5_1 and 5_2) was drawn in ChemDraw® and was subjected to molecular dynamic studies, giving different conformers for each topology. For each of the conformers, a CD spectrum was predicted using semi-empirical molecular

modelling (ZINDO-RPA, 80 excitations) implemented in ArgusLab 4.0.1. The theoretical CD was reconstructed in Gabedit (Lorentzian lineshape) and compared to the experimental CD spectrum. The final result takes into consideration two limitations: 1) the predicted CD needs to have a high degree of similarity with the experimental one, and 2) the computed conformer needs to be arranged in an approximate C_2 symmetry fashion. This idea is expressed in the diagram below (Scheme 4.2):



Scheme 4.2. The check diagram for determining the conformation of ***RSSR*-NBN knot**.

After a large number of CD spectra were predicted (I have done 40 simulations for every type of knot and each simulated CD – reconstructed using Gabedit), careful inspection and comparison with the experimental CD spectrum were performed. Firstly, the theoretical CD shape was compared visually with the experimental one, and the ones that were similar enough were compared using an estimation of the closeness of the fit using the Root Mean Square Error (RMSE) method (the equation can be found in the Experimental section (Chapter 7)) – the table with the similarity values can be found at the end of the Chapter (Table 4.1). We have concluded that the molecule formed is a 4_1 knot (Figure 4.21). This topology is also consistent with the NMR analysis: the molecule has a C_2 symmetry with two types of BDT cores, each core having different aromatic protons. The BDT and NDI cores are either far apart or in a T-shape position, explaining why there are not many cross peaks between the

NDI protons. Figure 4.22 shows the symmetrical parts of the knot, suggesting its C_2 symmetry.

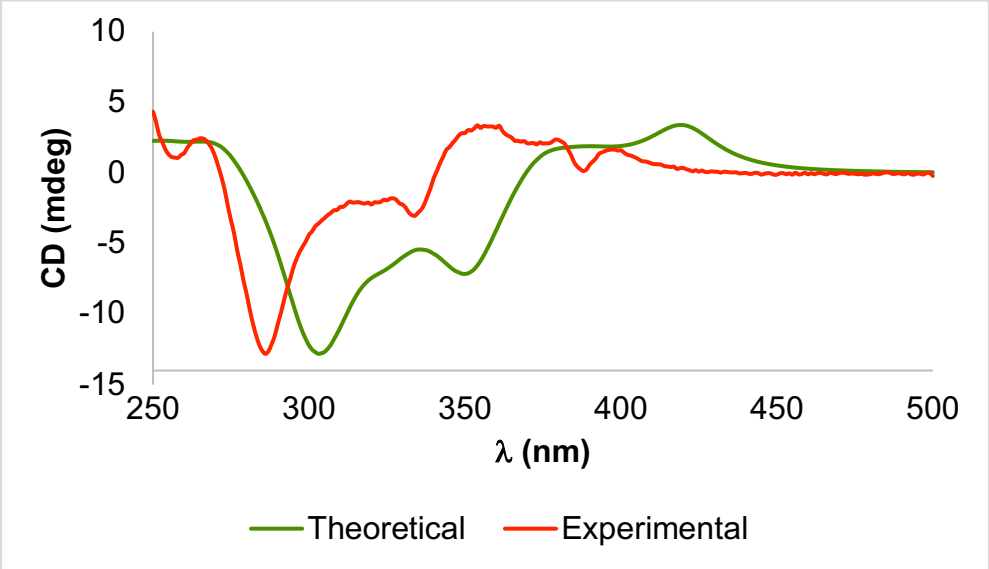


Figure 4.21. Superimposed theoretical (green) and experimental (red) CD spectra of the ***RSSR*-NBN knot**.

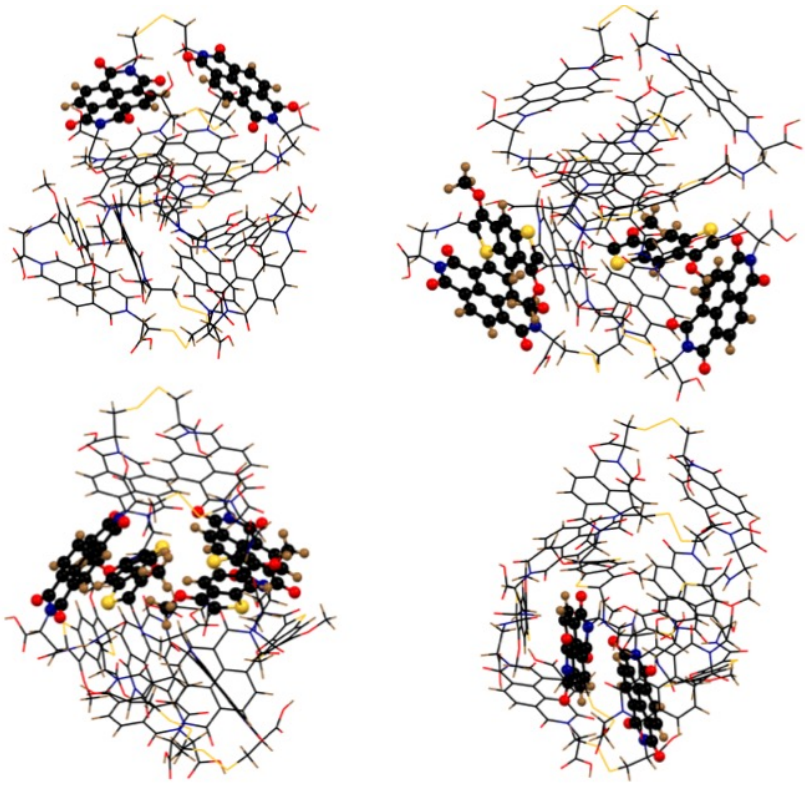


Figure 4.22. Simulated structure of ***RSSR*-NBN knot** using Avogadro in different arrangements (C – black, H – brown, O – red, N – blue, S – yellow).

Further understanding of the system can be obtained by varying the experimental conditions for DCL set up such as concentration of the building block. Therefore, DCLs of the trimer at different concentrations (0.5, 1, 2.5 and 5 mM) in the presence of 1 M $NaNO_3$ were set up. As the concentration is decreased, the percentage of knot decreased, but not significantly (Figure 4.23). The distribution of libraries of the trimer without any additives or salts at the same concentration range was also studied. For the DCL with no salt (Figure 4.24), the amount of the knot drops as the concentration of the trimer is decreased (even disappearing at 0.5 mM), whereas the formation of the monomer and the interlocked dimer is favoured. The interlocked dimer was isolated and analysed by MS; zoom scan and MS/MS suggesting it is a [2]catenane (Figure 4.25).

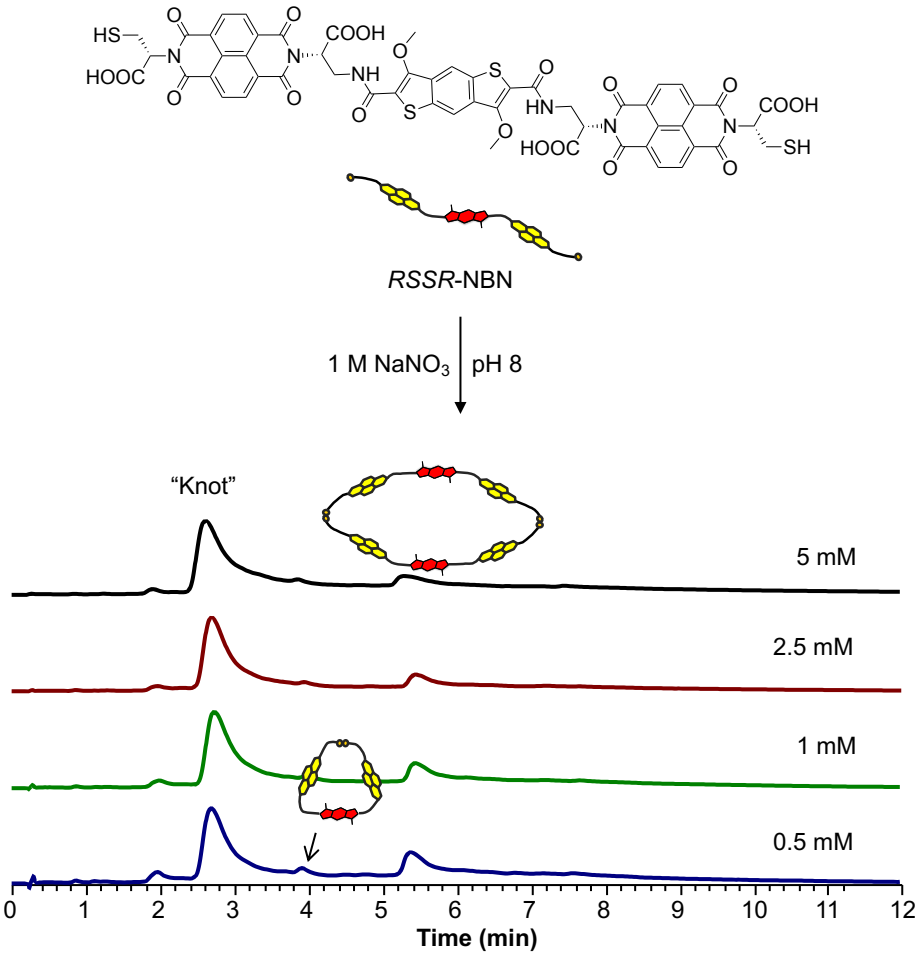


Figure 4.23. Reverse-phase HPLC analysis of ***RSSR*-NBN** library at 5, 2.5, 1 and 0.5 mM in the presence of 1 M $NaNO_3$. The unlabelled peaks did not ionise and could not be identified.

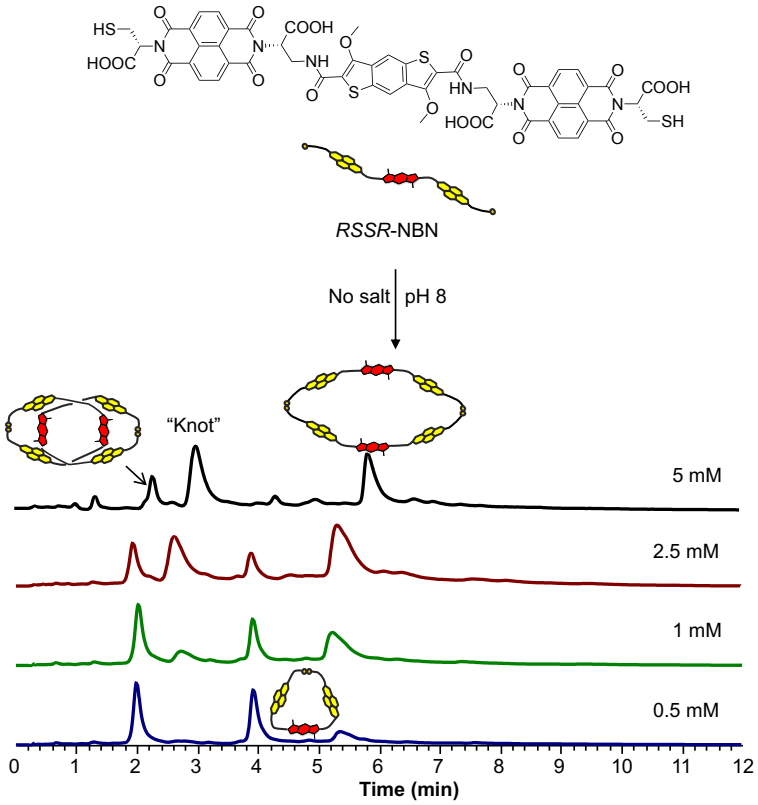


Figure 4.24. Reverse-phase HPLC analysis of ***RSSR*-NBN** library at 5, 2.5, 1 and 0.5 mM with no salt. The unlabelled peaks did not ionise and could not be identified.

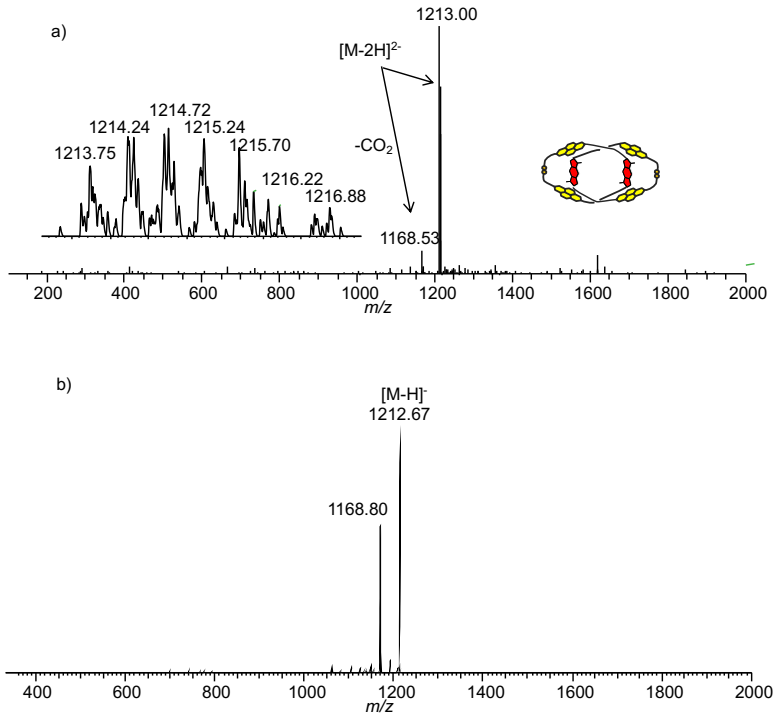


Figure 4.25. a) MS (*-ve*) of ***RSSR*-NBN [2]catenane**; zoom of the molecular ion is shown as an inset and b) MS/MS (*-ve*) of ***RSSR*-NBN [2]catenane**.

4.1.2. Analysis of *RSSR*-NBN [2]catenane

The chromatogram of the isolated [2]catenane is shown in Figure 4.26. The ^{1}H NMR spectrum (Figure 4.27) of the sample in D_2O looks simple, but interesting. The aromatic region (between 9 and 8.5 ppm) displays four doublets, which correspond to the NDI protons. The BDT protons appear more shielded compared to the NDI ones (5.61 ppm), suggesting that the BDT cores are sandwiched between the NDI units. The COSY spectrum (Figure 4.29) confirms that the doublets in the aromatic region are correlated, and also cross peaks between the aliphatic α and β protons are present. The NOESY spectrum (Figure 4.29) shows cross peaks between the NDI and BDT cores, and also between NDI and the methoxy group as well as between NDI and the aliphatic protons. The HSQC spectrum (Figure 4.30) further confirms the catenated structure; the ^{13}C NMR spectrum matches the expected pattern for a [2]catenane. The VT NMR (between 283 and 353 K; shown in Figure 4.28) indicated the structure is rigid: the peaks only become broad (because of the loss in definition) and shift upfield as the temperature is increased. CD and UV-Vis spectra of the isolated [2]catenane were also acquired and are shown in Figure 4.31. The absorbance spectrum displays the typical pattern of NDI scaffold centred around 380 nm and the band corresponding to the BDT appears hypsochromically shifted related to this. The CD spectrum is characterised by a small negative Cotton effect corresponding to the NDI-BDT (λ = 390 nm) band.

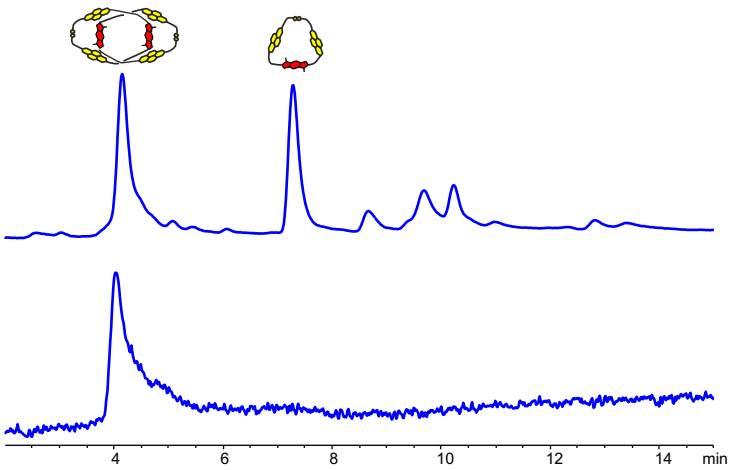


Figure 4.26. Reverse-phase HPLC analysis of ***RSSR*-NBN** (5 mM total concentration) library with no salt before (top trace) and ***RSSR*-NBN [2]catenane** after isolation (bottom trace). Absorbance spectra recorded at 389 nm. The unlabelled peaks did not ionise and could not be identified.

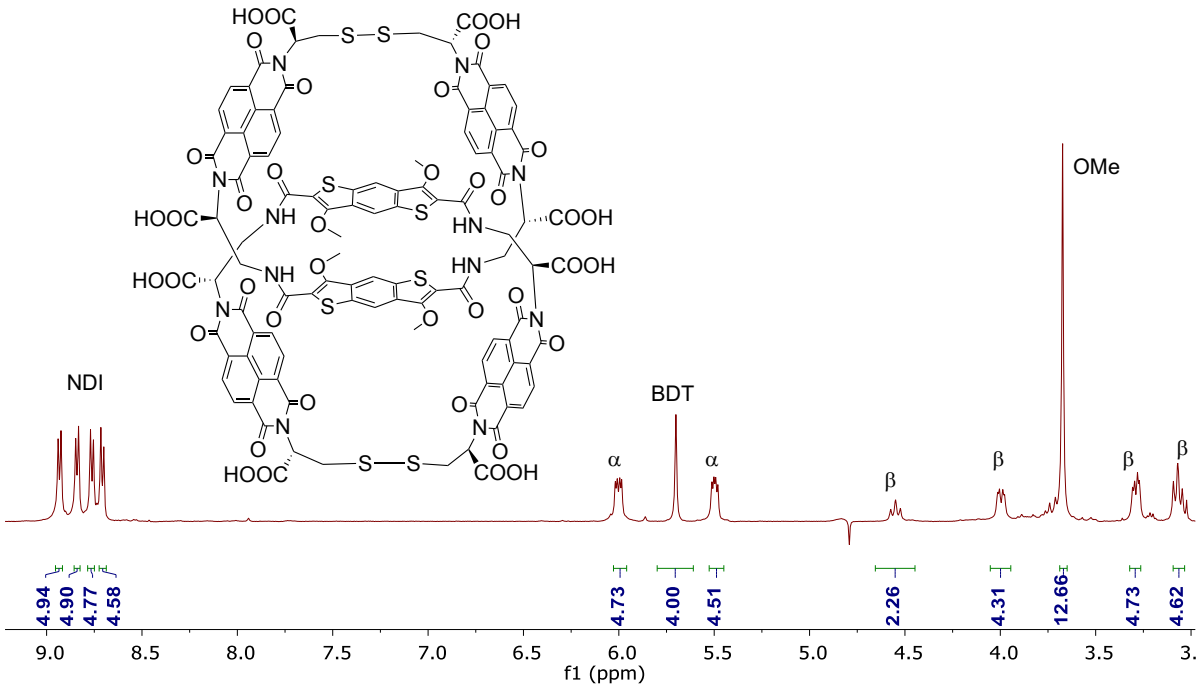


Figure 4.27. ^{1}H NMR spectrum (500 MHz, 298 K) of ***RSSR*-NBN [2]catenane**. The residual solvent (H_2O) was referenced at 4.79 ppm.

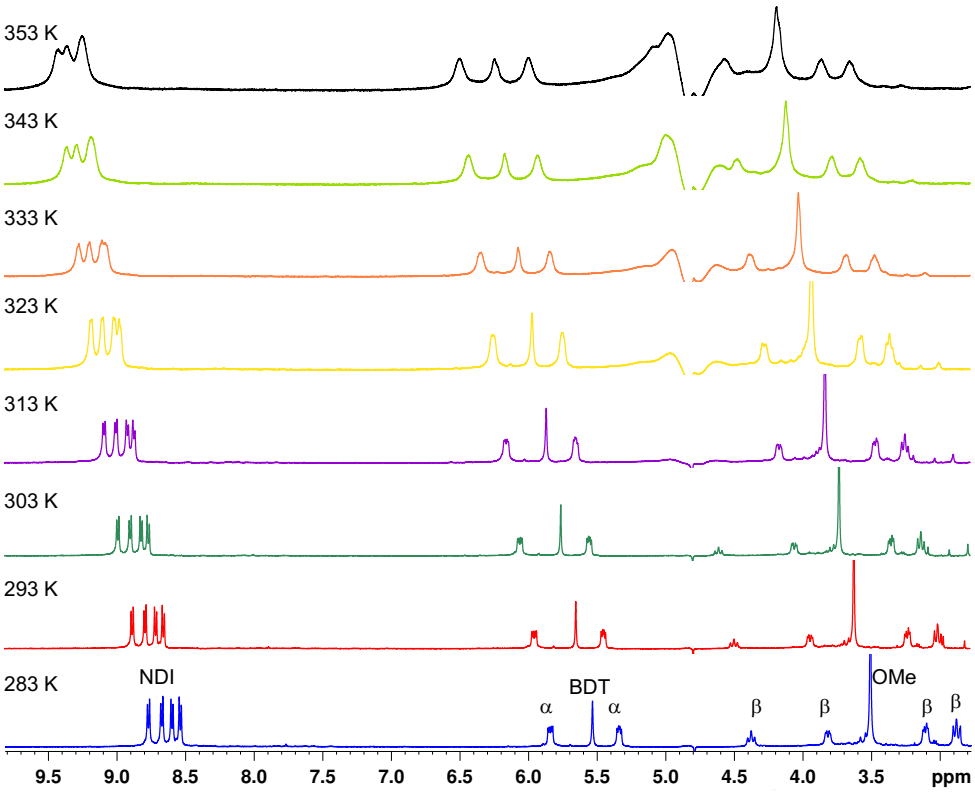


Figure 4.28. VT ^{1}H NMR spectra (500 MHz) of ***RSSR*-NBN [2]catenane** at the temperatures indicated above each spectrum. The residual solvent (H_2O) was referenced at 4.79 ppm.

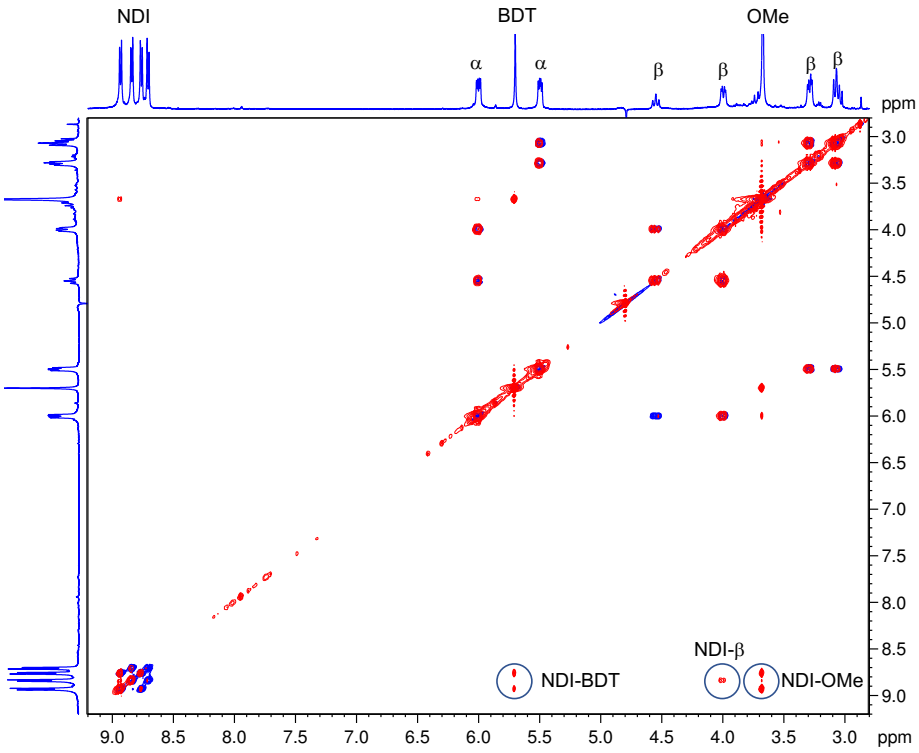


Figure 4.29. ^{1}H-^{1}H COSY (blue) and 2D NOESY (red) spectra superimposed of ***RSSR*-NBN [2]catenane**. The residual solvent (H_2O) was referenced at 4.79 ppm.

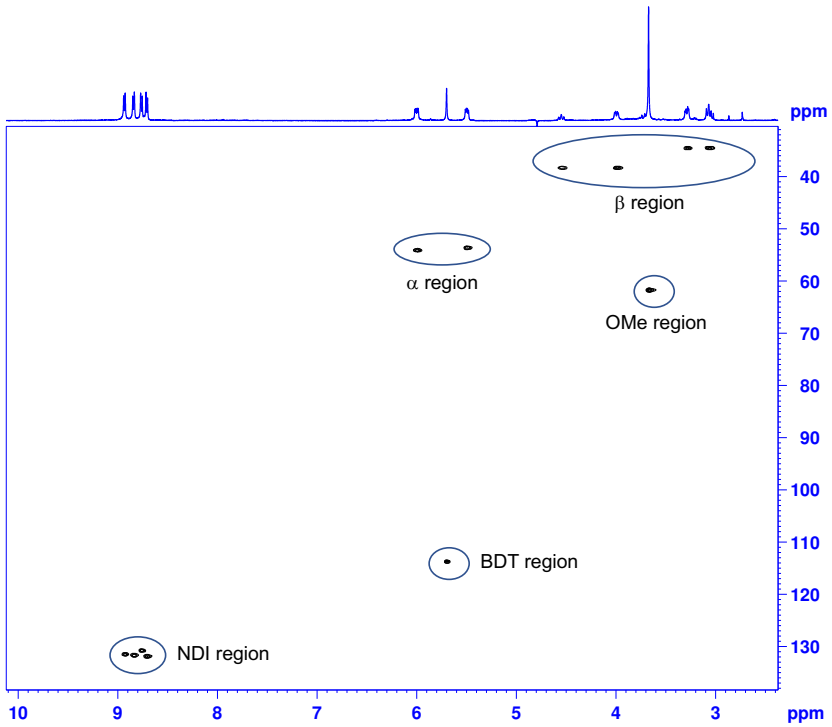


Figure 4.30. HSQC spectrum (500 MHz, 298 K) of ***RSSR*-NBN [2]catenane**. The residual solvent (H_2O) was referenced at 4.79 ppm.

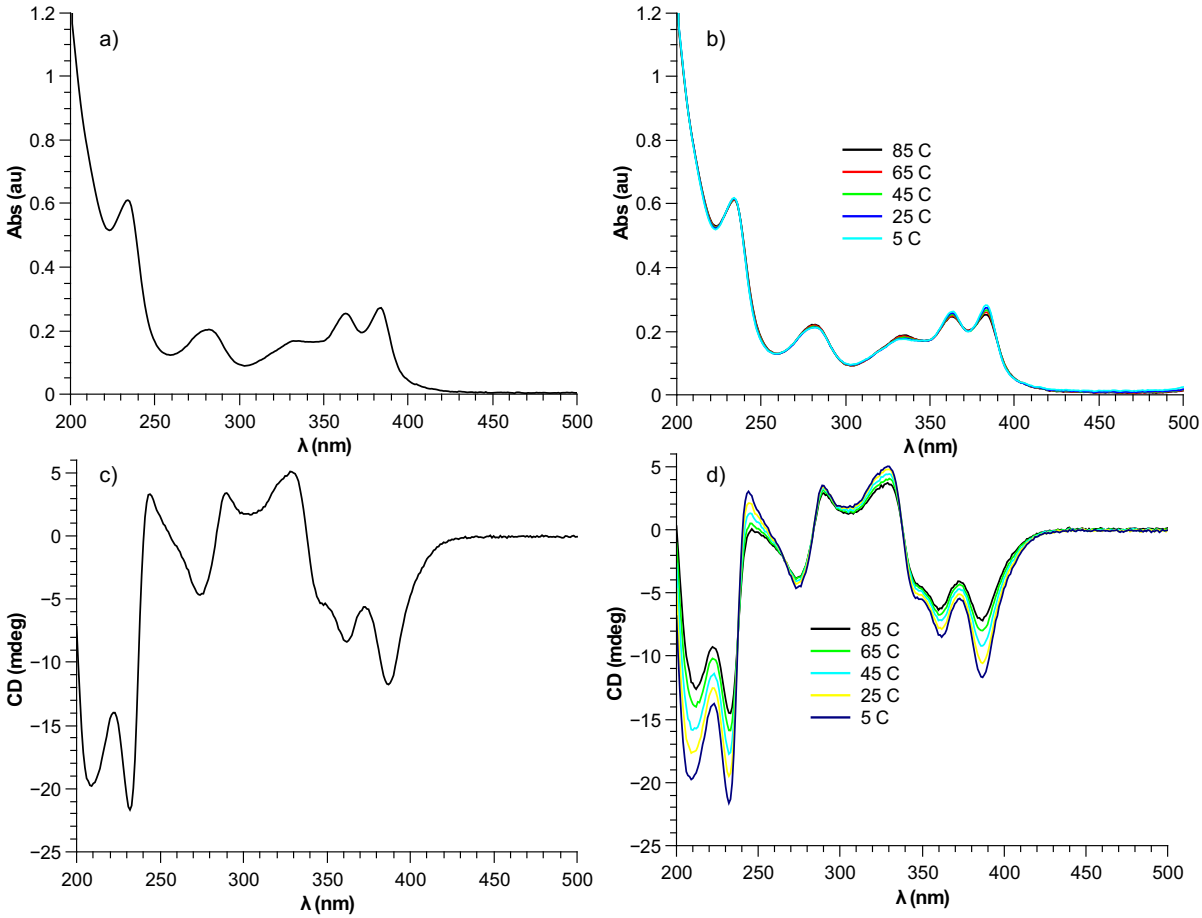
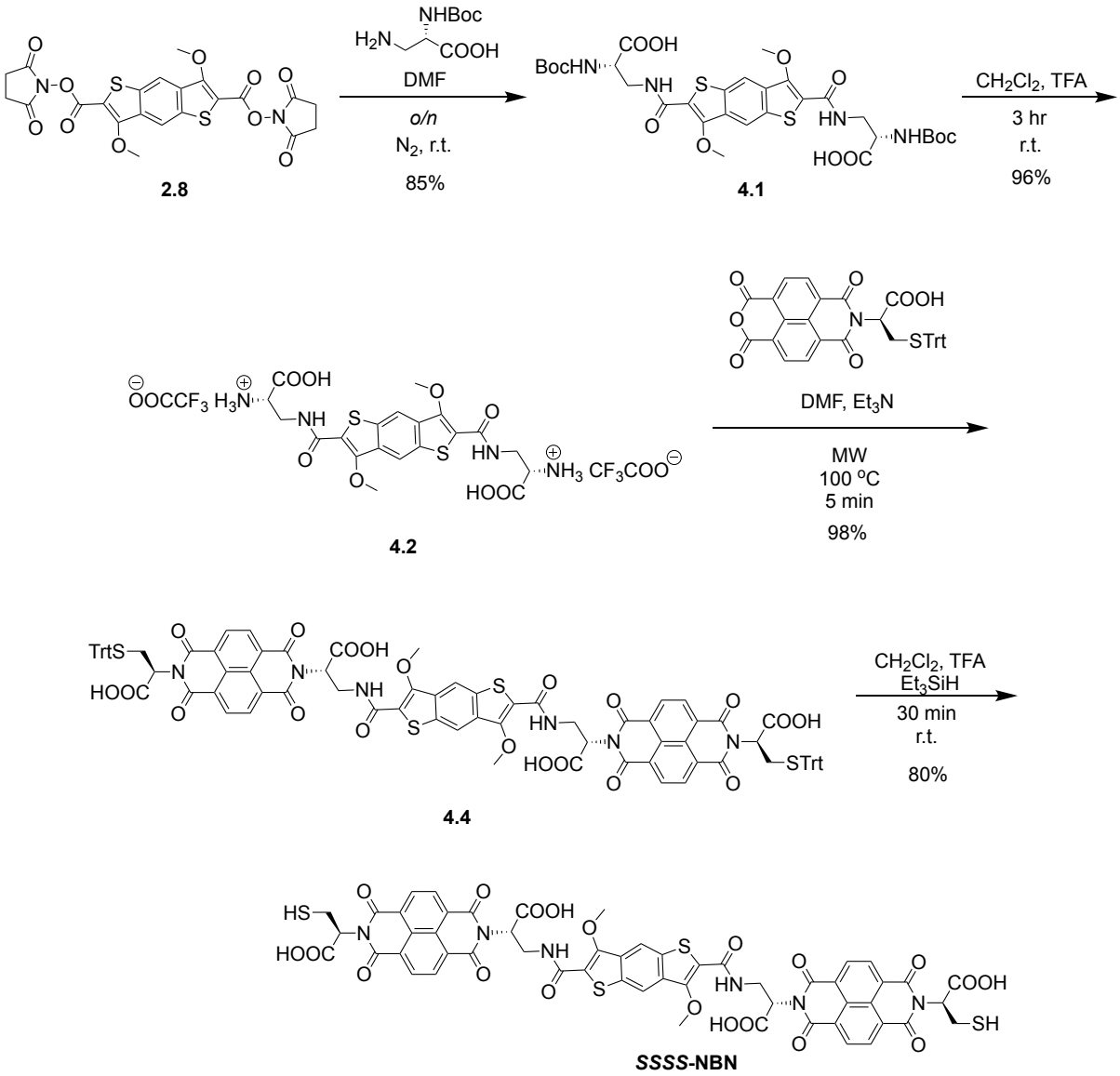


Figure 4.31. a) UV-Vis spectrum of ***RSSR*-NBN [2]catenane** at 23 °C (296 K), b) VT UV-Vis spectra of ***RSSR*-NBN [2]catenane** between 5 – 85 °C at the specified temperatures, c) CD spectrum of ***RSSR*-NBN [2]catenane** at 23 °C and d) VT CD spectra of ***RSSR*-NBN [2]catenane**, between 5 – 85 °C at the specified temperatures.

We have already showed the importance of chiral recognition in DCLs outcome in the previous chapters, thus this was also investigated for this system. The heterochiral trimer was therefore synthesised; synthetic route is depicted in Scheme 4.3. It resembles the synthesis of ***RSSR*-NBN**, the only difference being the reaction of **4.2** with NMI-Trt-*S*-Cys, followed by trityl deprotection to yield the desired product.



Scheme 4.3. The synthetic scheme for ***SSSS*-NBN**.

4.2. Analysis of *SSSS*-NBN library

DCLs were prepared with and without additives by dissolving the compound in NaOH 10 mM and adjusting the pH to 8 using 100 mM NaOH / 100 mM HCl. The distribution of the DCL with no salt or additives showed a monomer and a [2]catenane based on MS and MS/MS analyses (other peaks were present, but they did not ionise). In the presence of 1 M $NaNO_3$ (which increases the hydrophobic effect), the major component is the [2]catenane (60%). When the DCL is prepared with 1 M $TBANO_3$, which solvates the NDI core, the species forming most is a monomer – Figure 4.32. Chirality strongly influences the distribution of the DCL: if the homochiral trimer

assembles into a knot, the heterochiral one makes a [2]catenane as major components in the presence of salt (Figures 4.33 and 4.34). These results strengthen further the importance of chiral recognition and mismatch in the synthesis of interlocked molecules through DCC, which is also observed in biological processes in Nature.

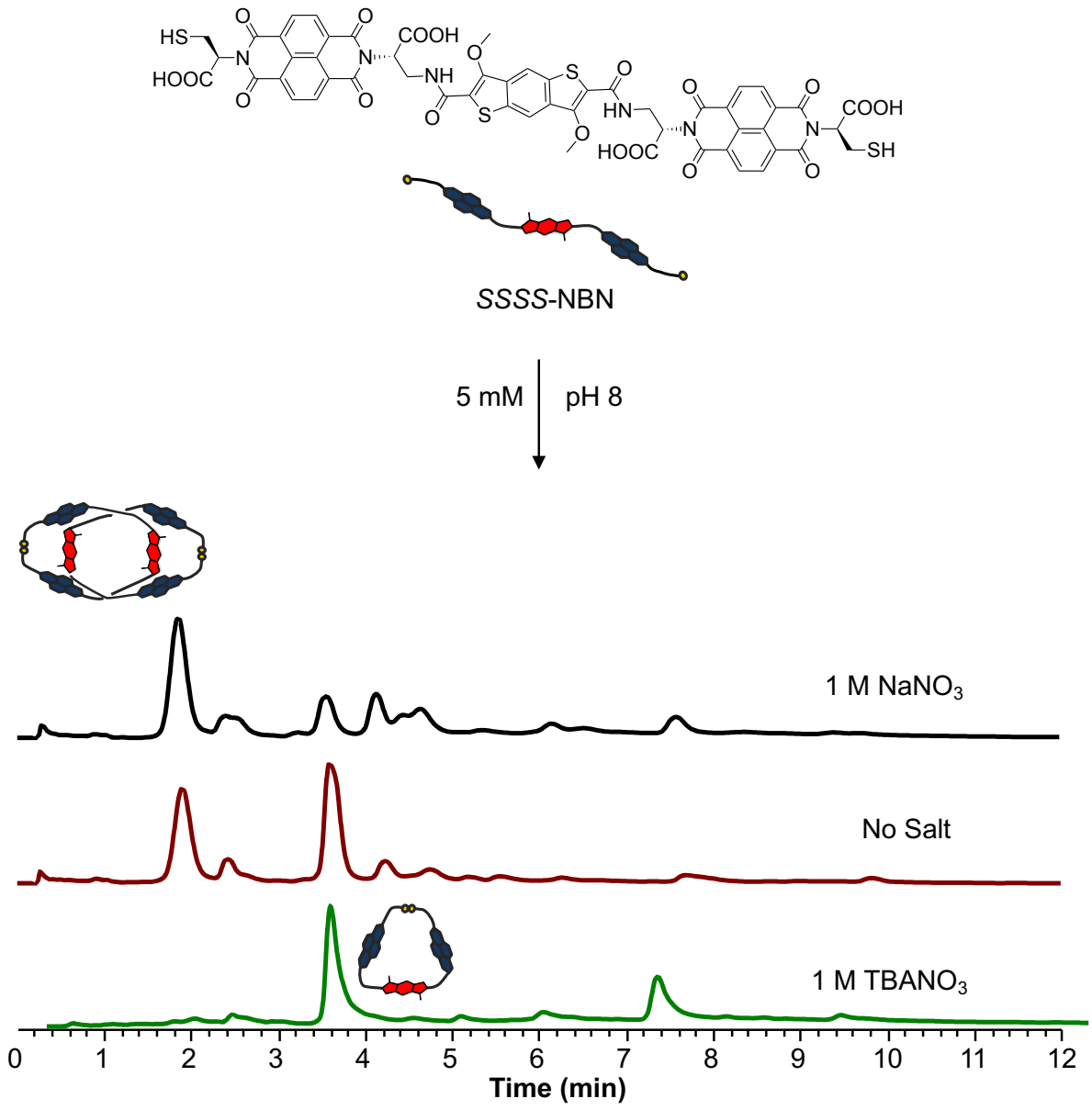


Figure 4.32. Reverse-phase HPLC analysis of ***SSSS*-NBN** (5 mM total concentration) library in the presence of 1 M $NaNO_3$, no salt, and in the presence of 1 M $TBANO_3$. Absorbances recorded at 389 nm. The unlabelled peaks did not ionise and could not be identified.

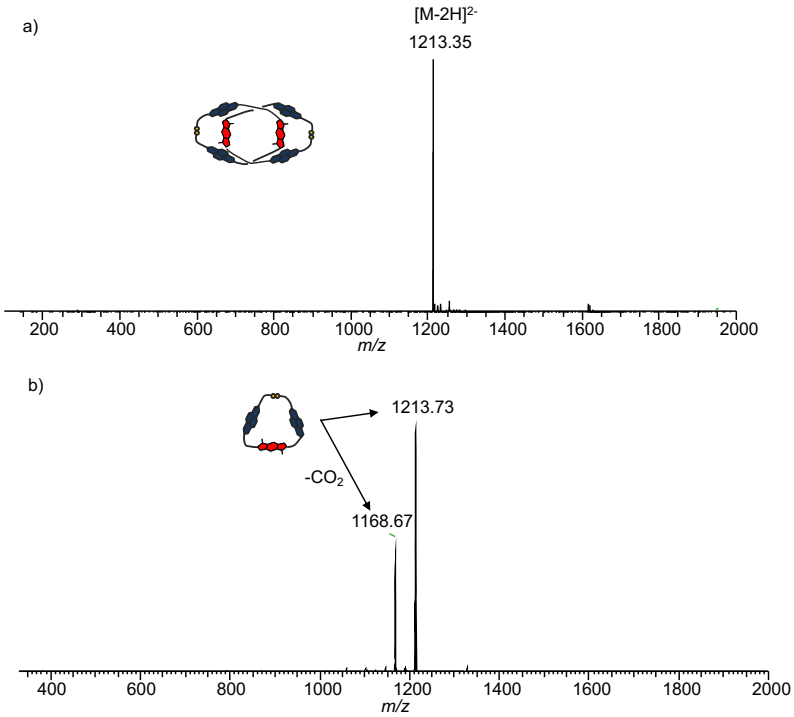


Figure 4.33. a) MS (-*ve*) of ***SSSS*-NBN [2]catenane** and b) MS/MS (-*ve*) of ***SSSS*-NBN [2]catenane**.

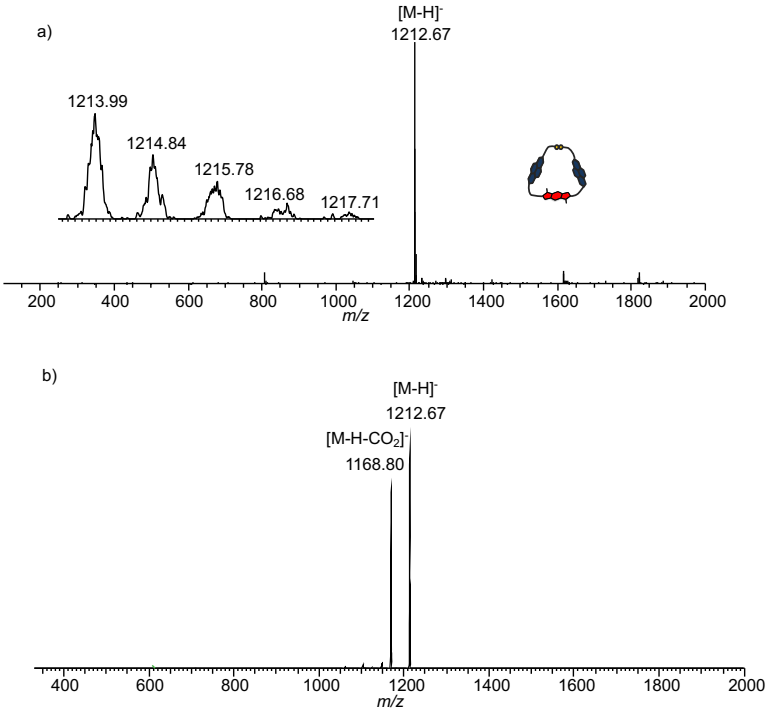


Figure 4.34. a) MS (-*ve*) of ***SSSS*-NBN monomer**; zoom of the molecular ion is shown as an inset and b) MS/MS (-*ve*) of ***SSSS*-NBN monomer**.

DCLs with ***SSSS*-NBN** at different concentrations were also prepared. In both cases (with and without $NaNO_3$) and as the concentration decreases, the amount of [2]catenane drops, while that of the monomer increases according to the Le Chatelier's principle (Figures 4.35 and 4.36).

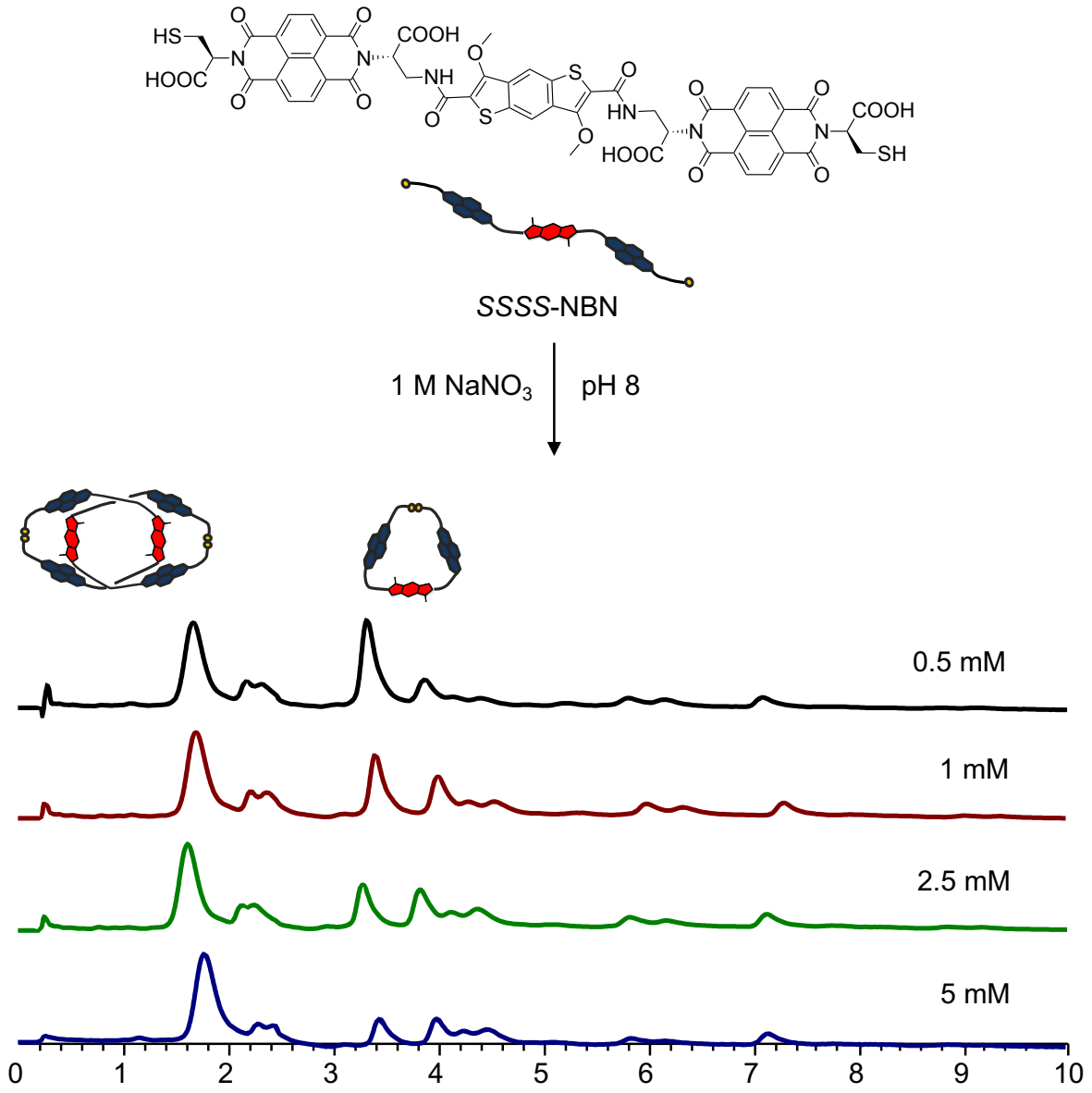


Figure 4.35. Reverse-phase HPLC analysis of ***SSSS*-NBN** library at 5, 2.5, 1 and 0.5 mM in the presence of 1 M $NaNO_3$. The unlabelled peaks did not ionise and could not be identified.

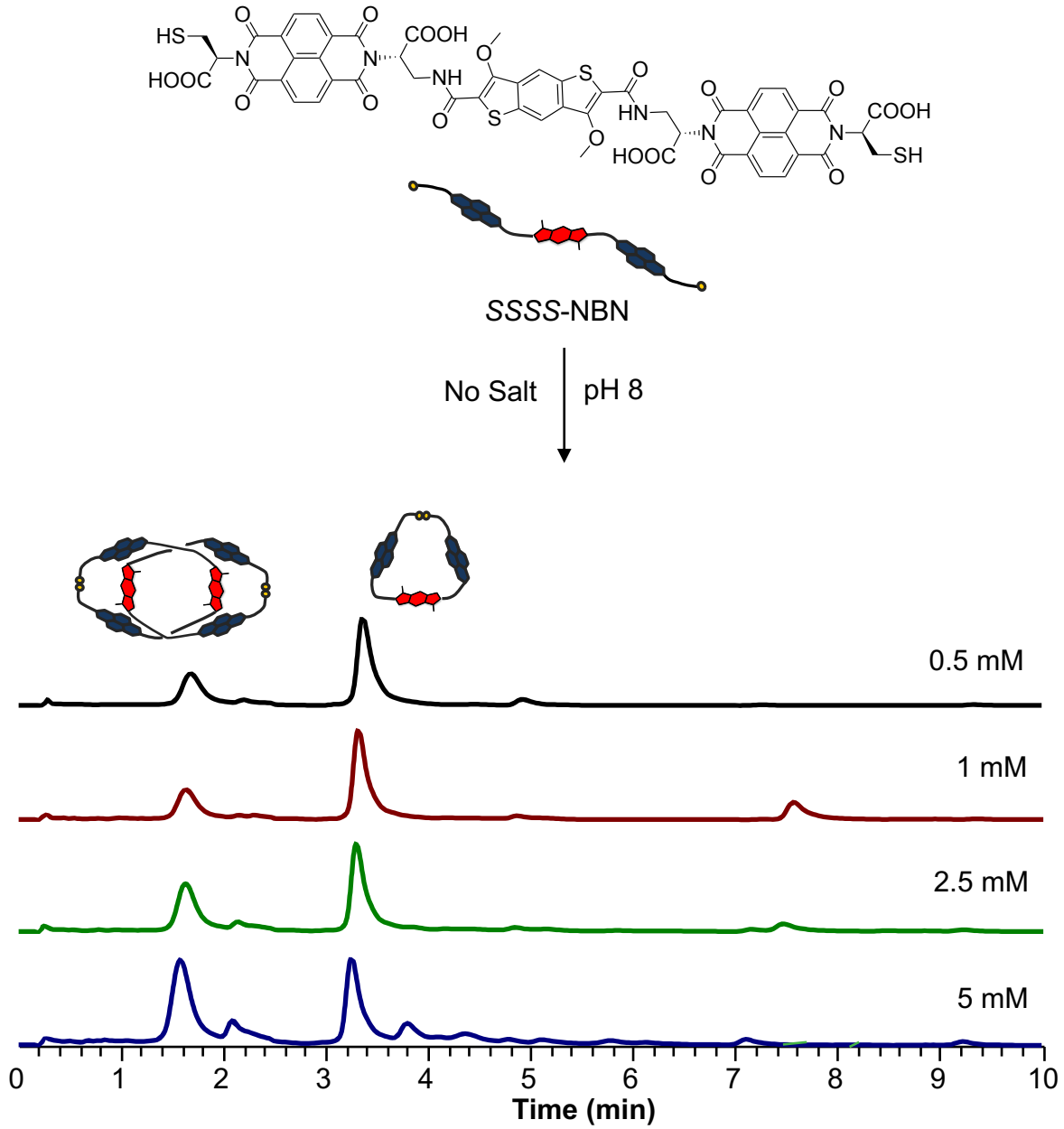


Figure 4.36. Reverse-phase HPLC analysis of ***SSSS*-NBN** library at 5, 2.5, 1 and 0.5 mM with no salt. The unlabelled peaks did not ionise and could not be identified.

4.2.1. Analysis of *SSSS*-NBN [2]catenane

The [2]catenane was isolated from a large library (> 3 mL) using semi-preparative HPLC and the purity of the isolated batch tested (Figure 4.37). The ^{1}H NMR spectrum shows two broad singlets in the NDI aromatic region, indicating a less rigid arrangement compared to its diastereomer (*i.e.* ***RSSR*-NBN [2]catenane**). The BDT peak is less shielded (6.14 ppm *vs.* 5.62 ppm in the diastereomeric

[2]catenane). This suggests that the molecule is less compact, so the BDT protons are less affected by the ring current of the other aromatics present in the molecule. The VT NMR spectra (Figure 4.38 – between 283 and 353 K) show that the peaks broaden, lose definition and move upfield as the temperature is increased, but the multiplicity is generally maintained the same throughout. At lower temperatures (283 and 293 K), one of the NDI peaks split into small broad singlets. At this temperature, there is less available energy for the system. The movement of the molecules is slow at the NMR timescale to distinguishing between different protons. NOESY spectrum does not offer extra information about the structure and arrangement (Figure 4.39). The HSQC spectrum confirms that the shifted singlet (6.14 ppm) corresponds to the BDT proton (Figure 4.40). The UV-Vis and CD profiles of this [2]catenane are similar to its diastereomer, with typical bands for NDI and BDT cores. The UV-Vis and CD VT studies indicate that the molecule is stable at high temperatures and does not undergo photodecomposition (due to the fact that initial and the final spectra were superimposable). This is in contrast with its diastereomer. (Figure 4.41).

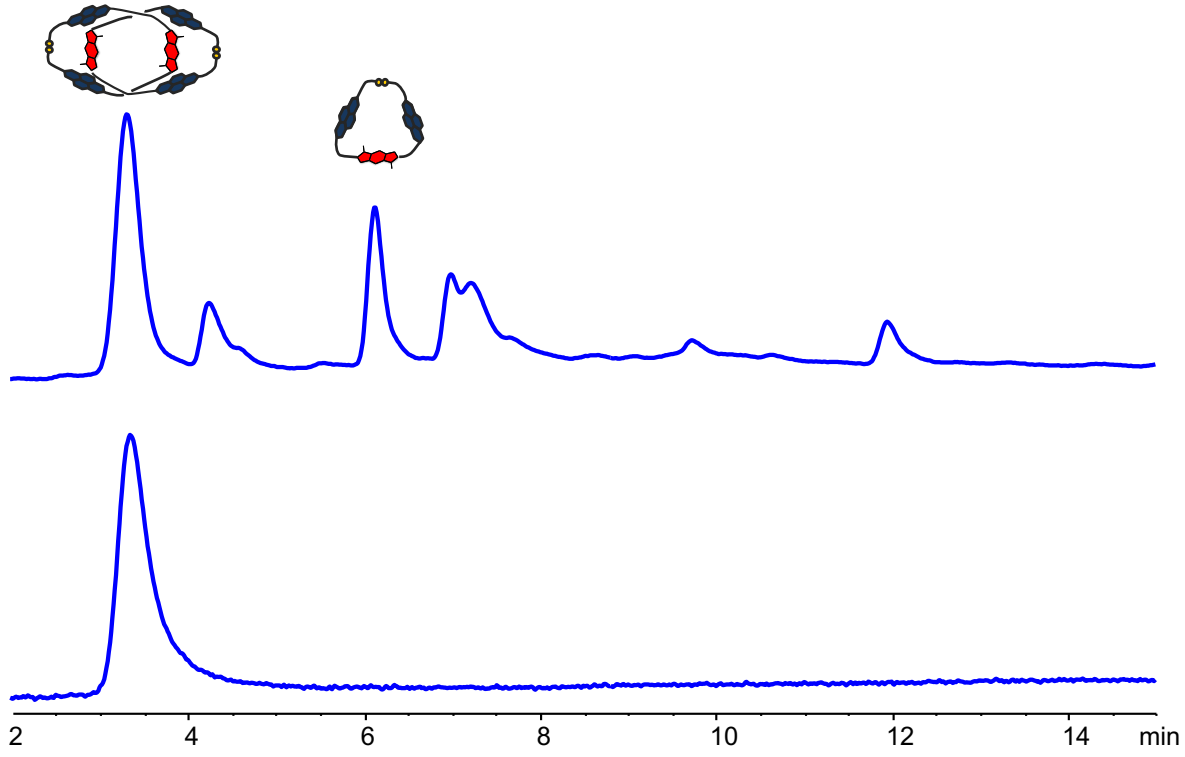


Figure 4.37. Reverse-phase HPLC analysis of ***SSSS*-NBN** (5 mM total concentration) library before (top trace) and ***SSSS*-NBN [2]catenane** after isolation (bottom trace). Absorbance spectra recorded at 389 nm. The unlabelled peaks did not ionise and could not be identified.

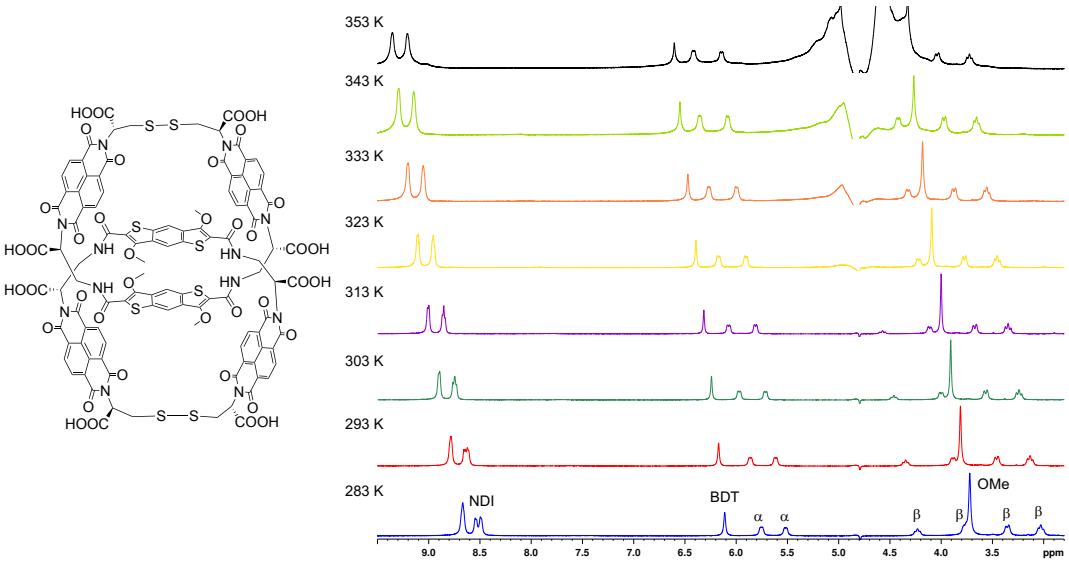


Figure 4.38. VT ^{1}H NMR spectra (500 MHz) of ***SSSS*-NBN [2]catenane** at the temperatures indicated above each spectrum. The residual solvent (H_2O) was referenced at 4.79 ppm.

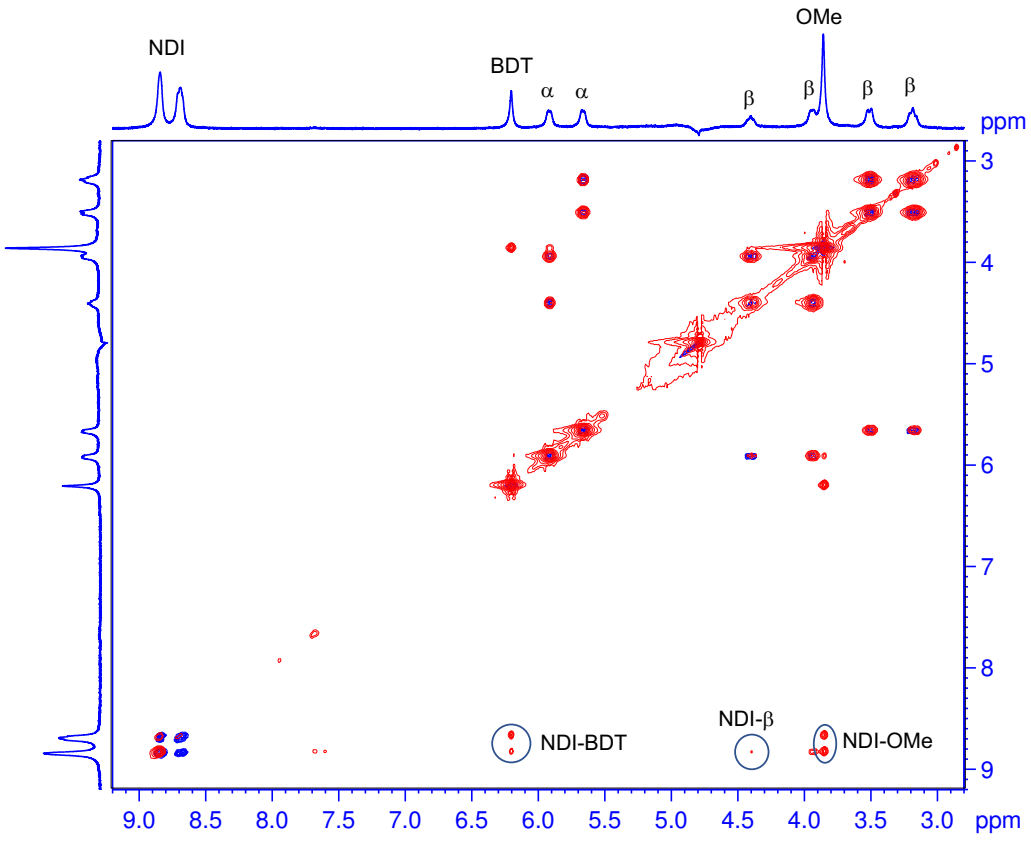


Figure 4.39. ^{1}H-^{1}H COSY (blue) and 2D NOESY (red) spectra (500 MHz, 298 K) superimposed of ***SSSS*-NBN [2]catenane**. The residual solvent (H_2O) was referenced at 4.79 ppm.

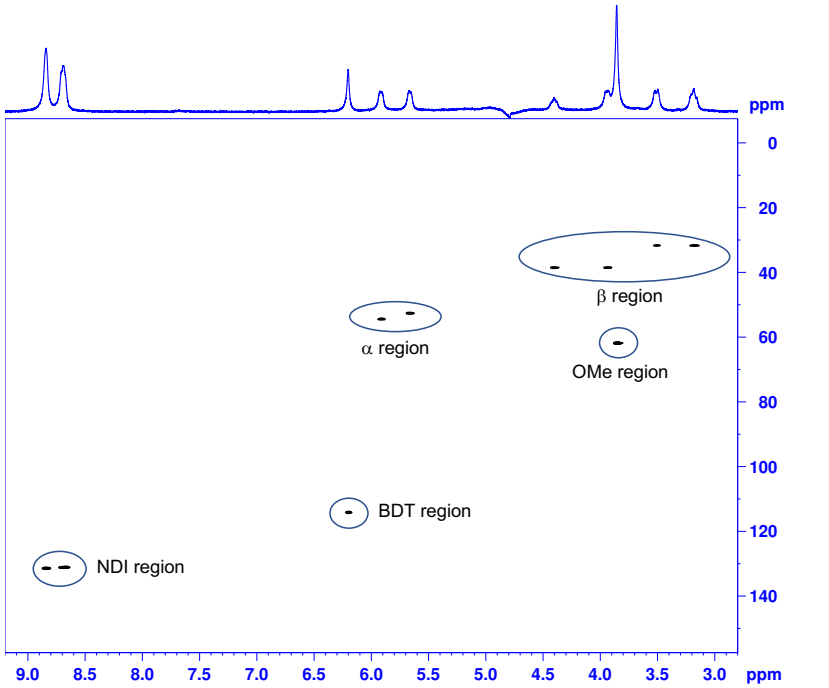


Figure 4.40. HSQC spectrum (500 MHz, 298 K) of ***SSSS*-NBN [2]catenane**. The residual solvent (H_2O) was referenced at 4.79 ppm.

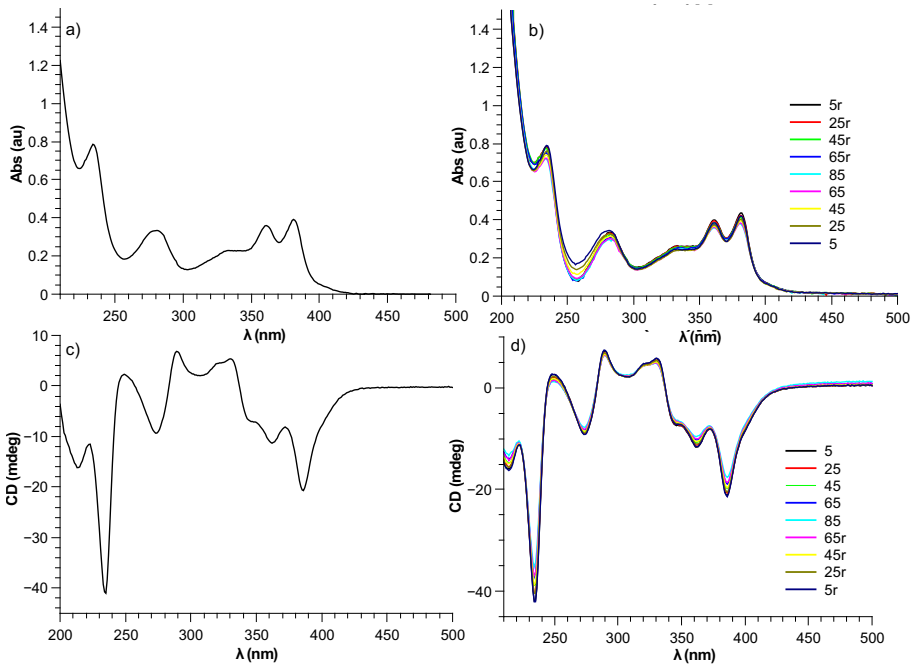
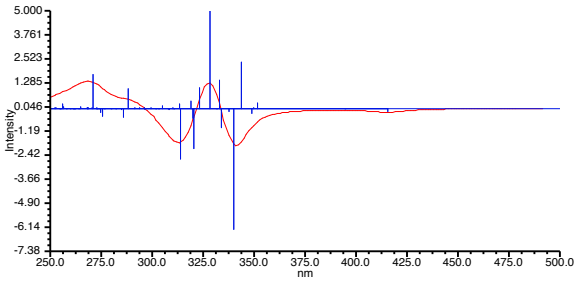


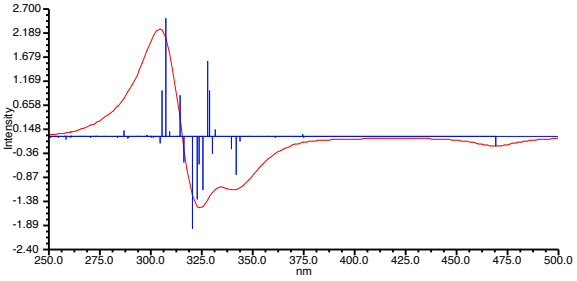
Figure 4.41. a) UV-Vis spectrum of ***SSSS*-NBN [2]catenane** at 23 °C (296 K), b) VT UV-Vis spectra of ***SSSS*-NBN [2]catenane** between 5 – 85 – 5 °C at the specified temperatures, c) CD spectrum of ***SSSS*-NBN [2]catenane** at 23 °C and d) VT CD spectra of ***SSSS*-NBN [2]catenane**, between 5 – 85 – 5 °C at the specified temperatures (r – return temperature).

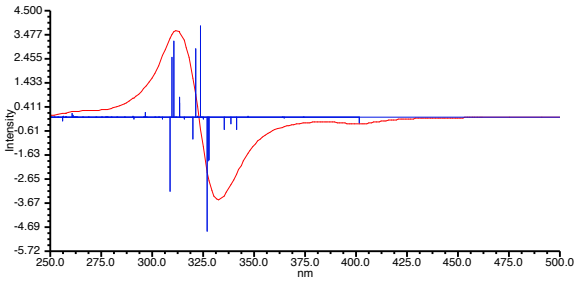
4.3 Conclusion

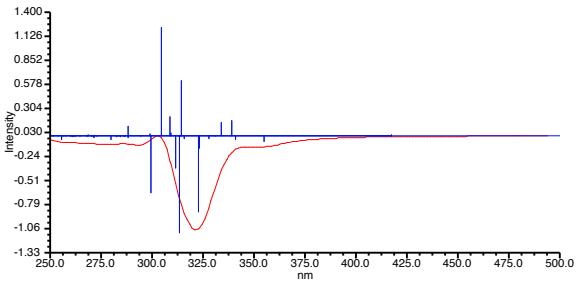
In conclusion, this chapter reports the synthesis of a 4_1 knot through DCC starting from a homochiral ***RSSR*-NBN** trimer. This work outlines the importance of building block design in DCL outcome; expanding the number of cores can lead to other interlocked molecules than [2]catenanes. 2D NMR spectroscopy has provided useful structural information, and the exact type of knot was determined based on the CD prediction – unfortunately even if over 400 crystallisations were set, none of them was successful in producing X-ray diffraction quality crystals. When the library concentration was decreased from 5 to 0.5 mM, the knot was not formed anymore, but instead a [2]catenane with a very compact structure was obtained (the BDT protons were shielded to 5.62 ppm). The effect of chirality was also studied; no knotted structure was formed in the heterochiral DCL even in the presence of salt, highlighting again the importance of chiral recognition in the synthesis of interlocked molecules through DCC. In this case, the DCL contains a [2]catenane as the major component, which is less rigid and compact in comparison to its diastereomer.

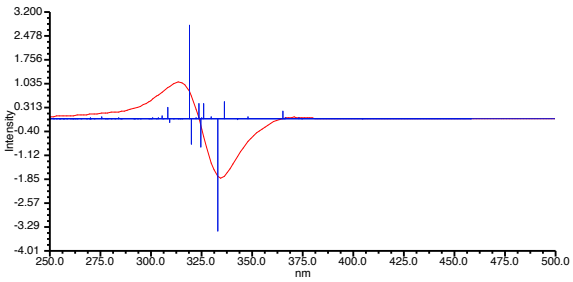
All CD predictions are given below (Figures 4.42 – 4.53)

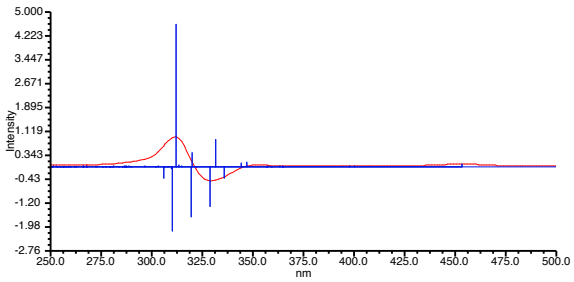

Intensity
nm

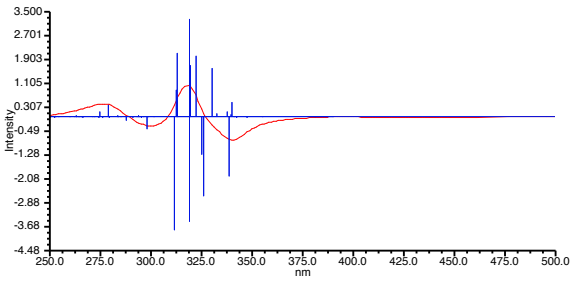

Intensity
nm

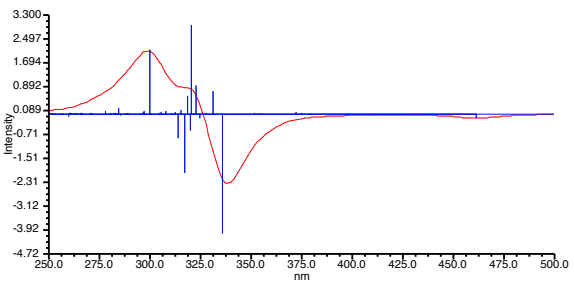

Intensity
nm

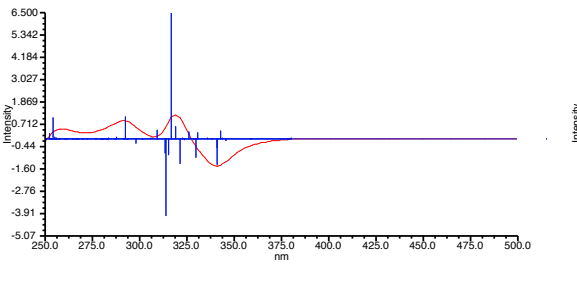

Intensity
nm

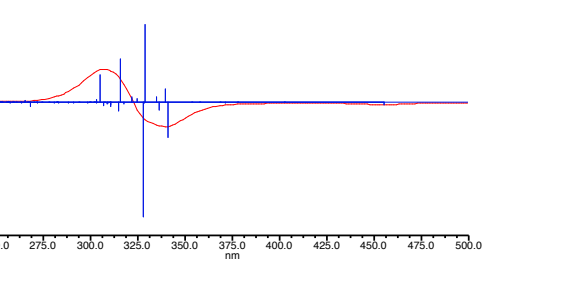

Intensity
nm

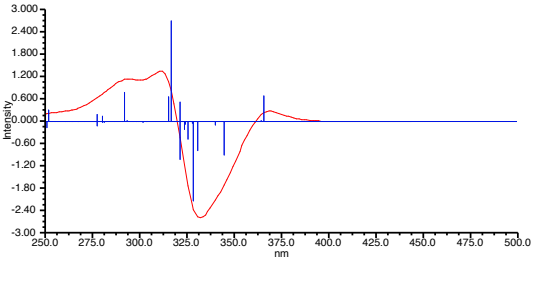

Intensity
nm

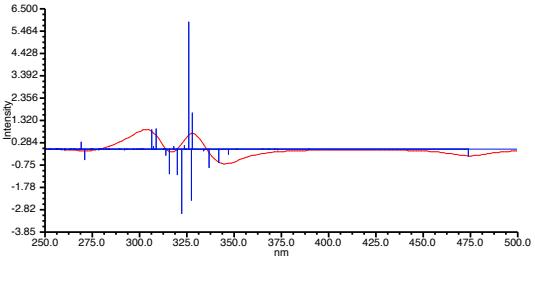

Intensity
nm


Intensity
nm


Intensity
nm


Intensity
nm


Intensity
nm


Intensity
nm

Intensity
nm


Intensity
nm

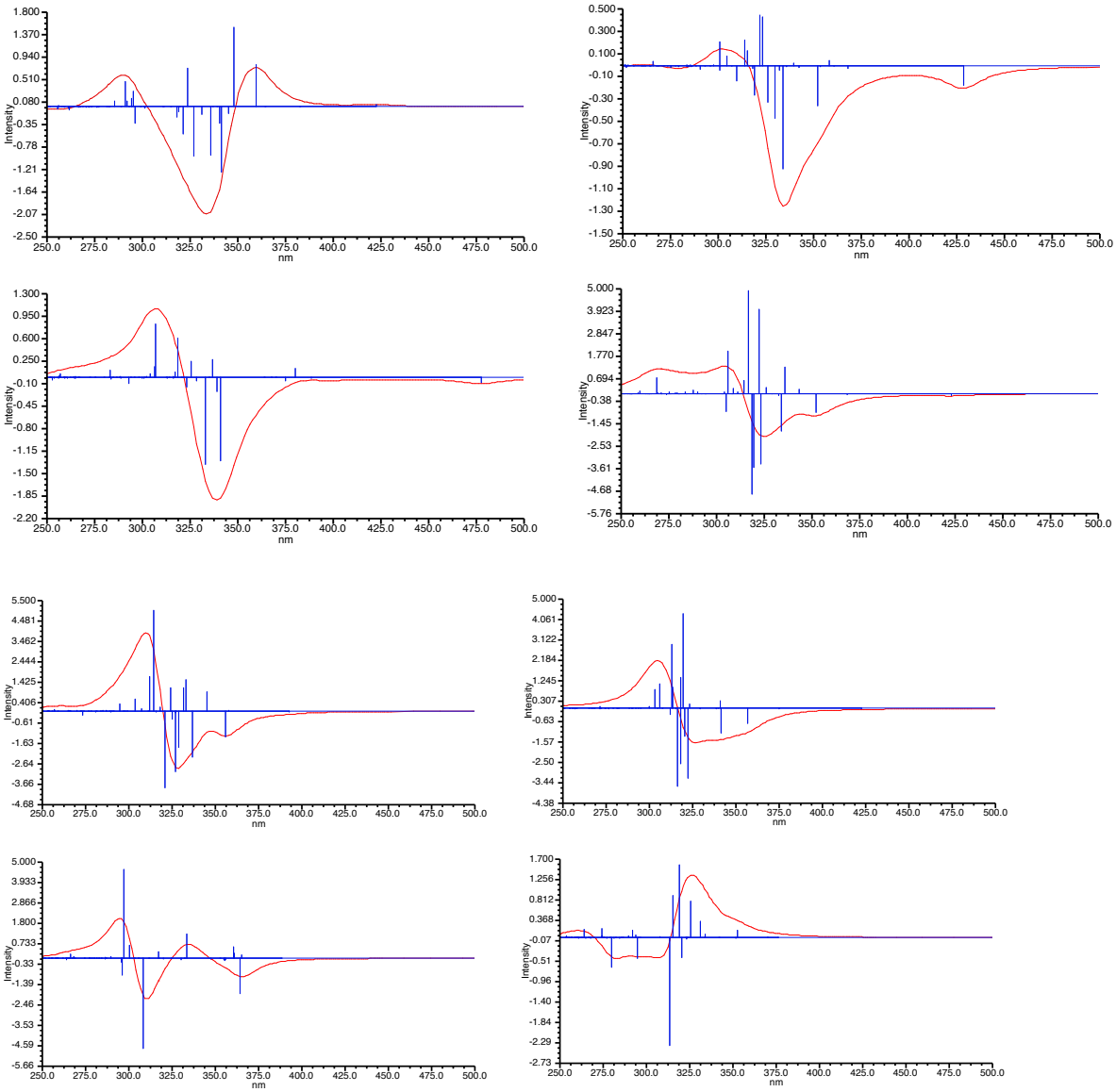


Figure 4.42. CD predictions of 3_1 knot using molecular dynamics *via* method a (described in Chapter 7).

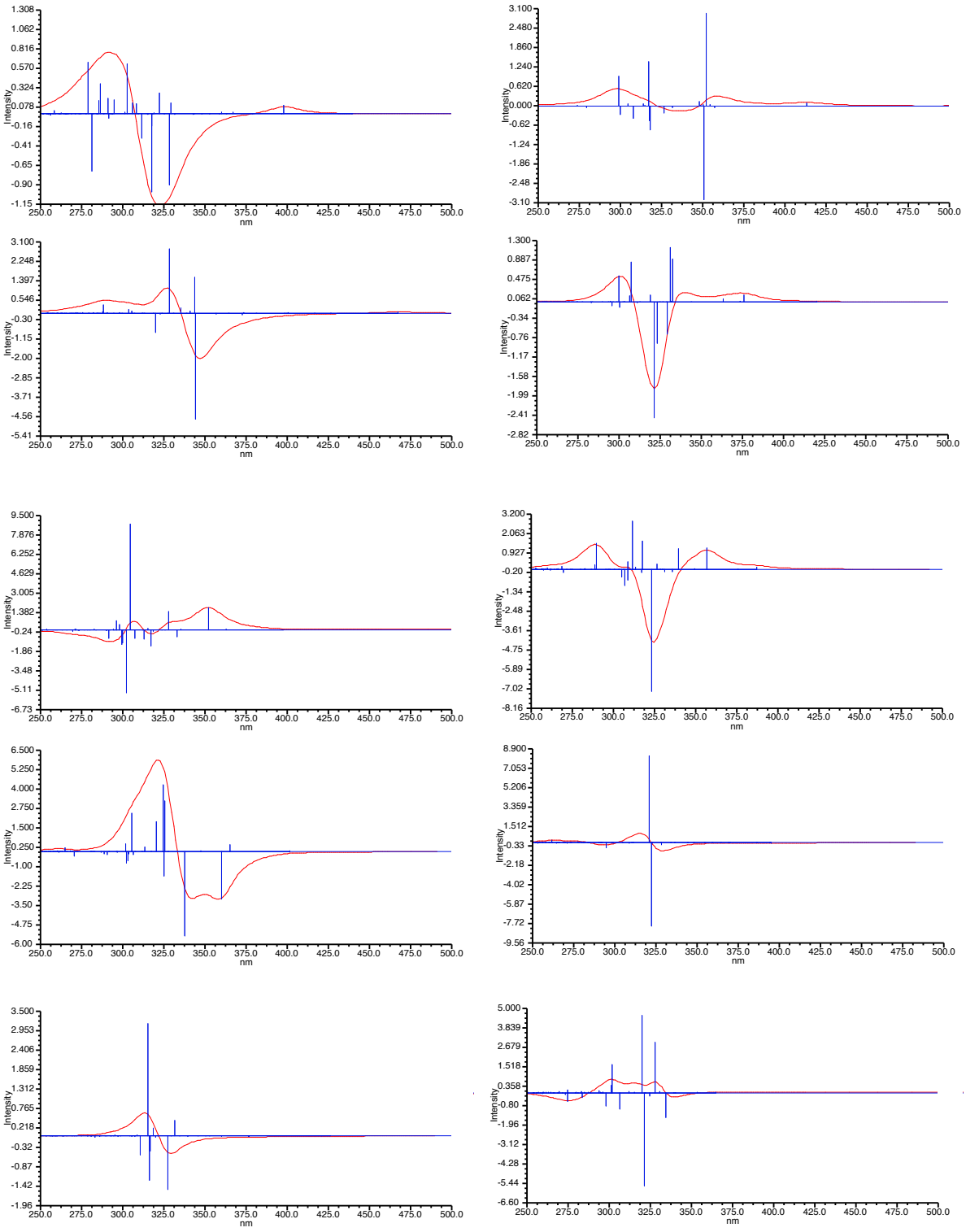


Figure 4.43. CD predictions of 3_1 knot using molecular dynamics *via* method b (described in Chapter 7).

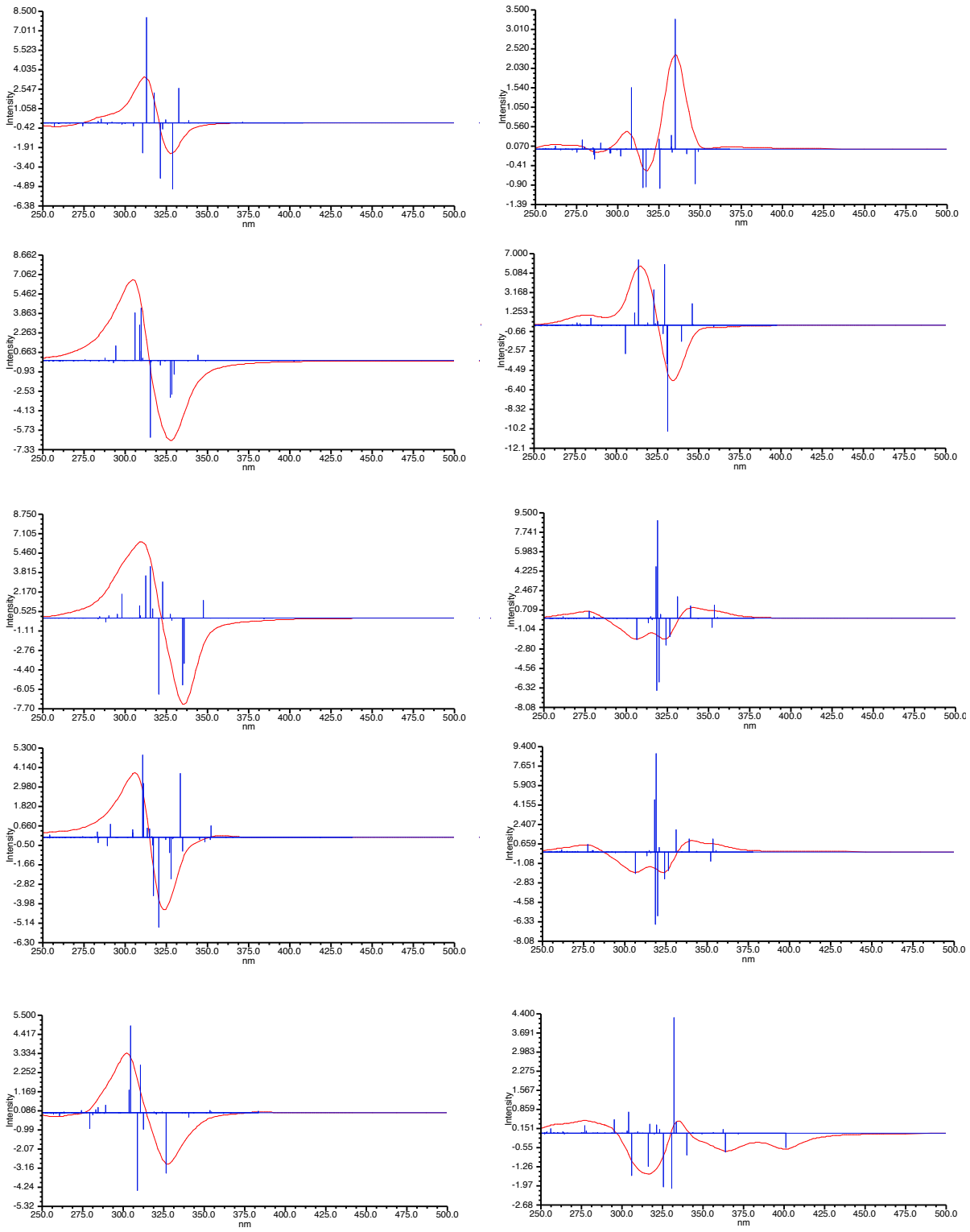
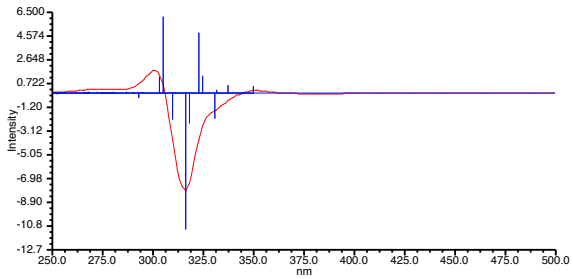
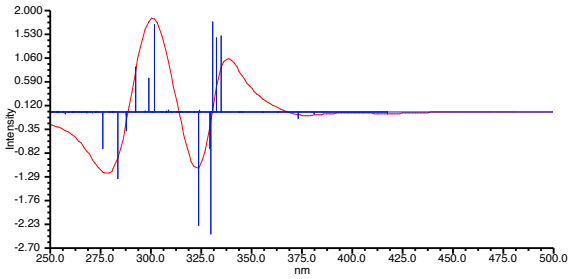
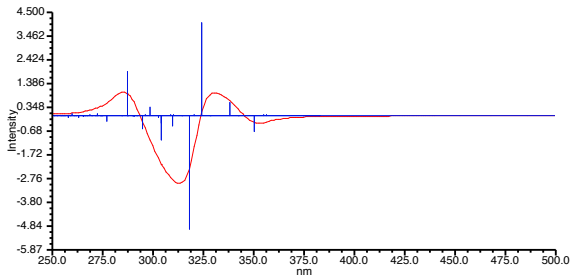
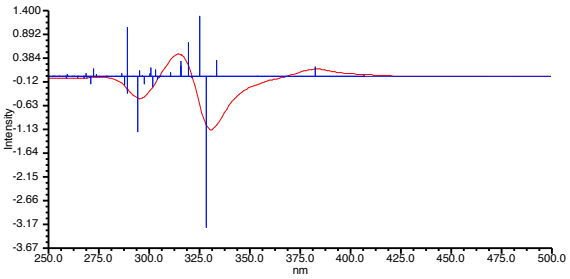


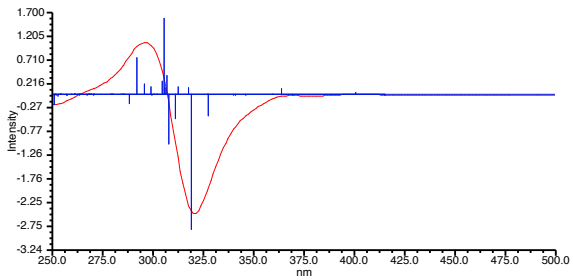
Figure 4.44. CD predictions of 3_1 knot using molecular dynamics *via* method c (described in Chapter 7).

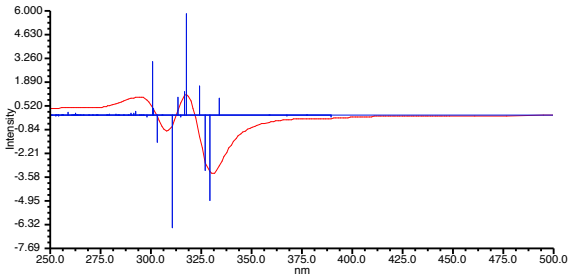

Intensity
nm

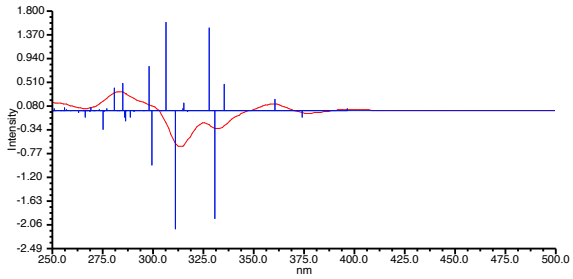

Intensity
nm

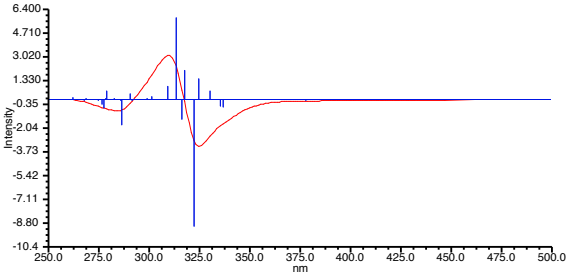

Intensity
nm

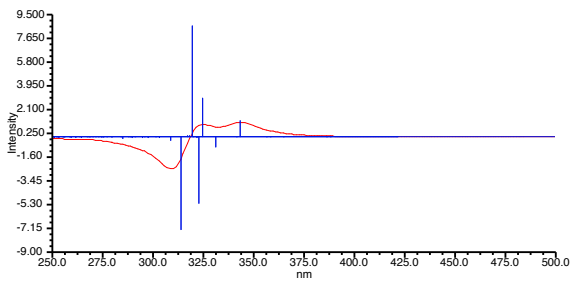

Intensity
nm

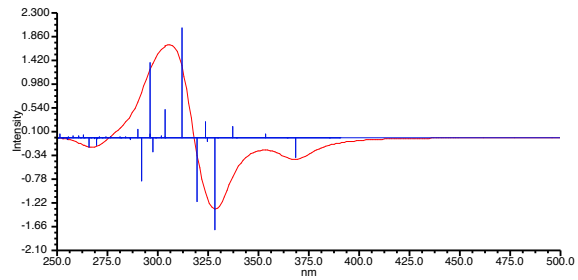

Intensity
nm

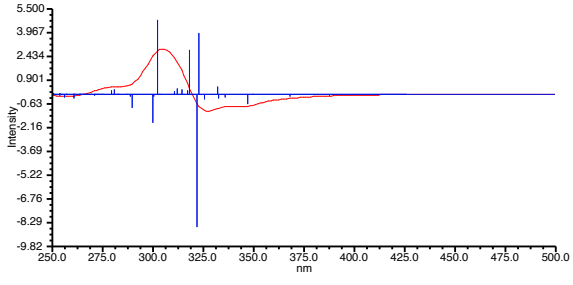

Intensity
nm

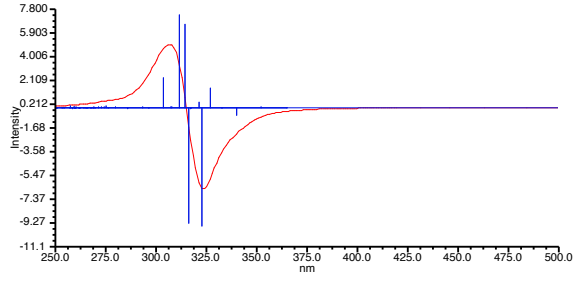

Intensity
nm


Intensity
nm


Intensity
nm


Intensity
nm


Intensity
nm


Intensity
nm

Intensity
nm

Intensity
nm

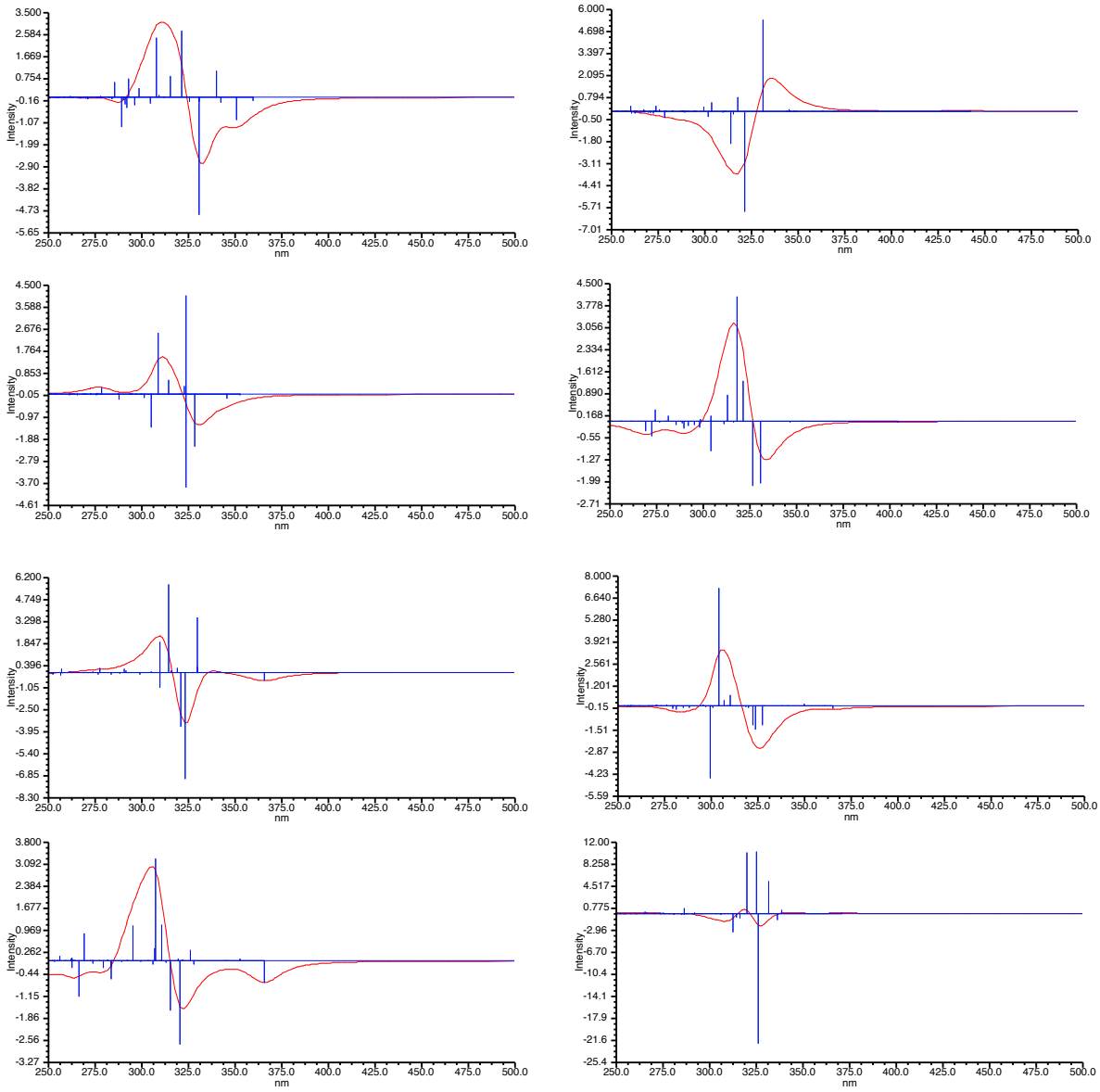


Figure 4.45. CD predictions of 4_1 knot using molecular dynamics *via* method a (described in Chapter 7).

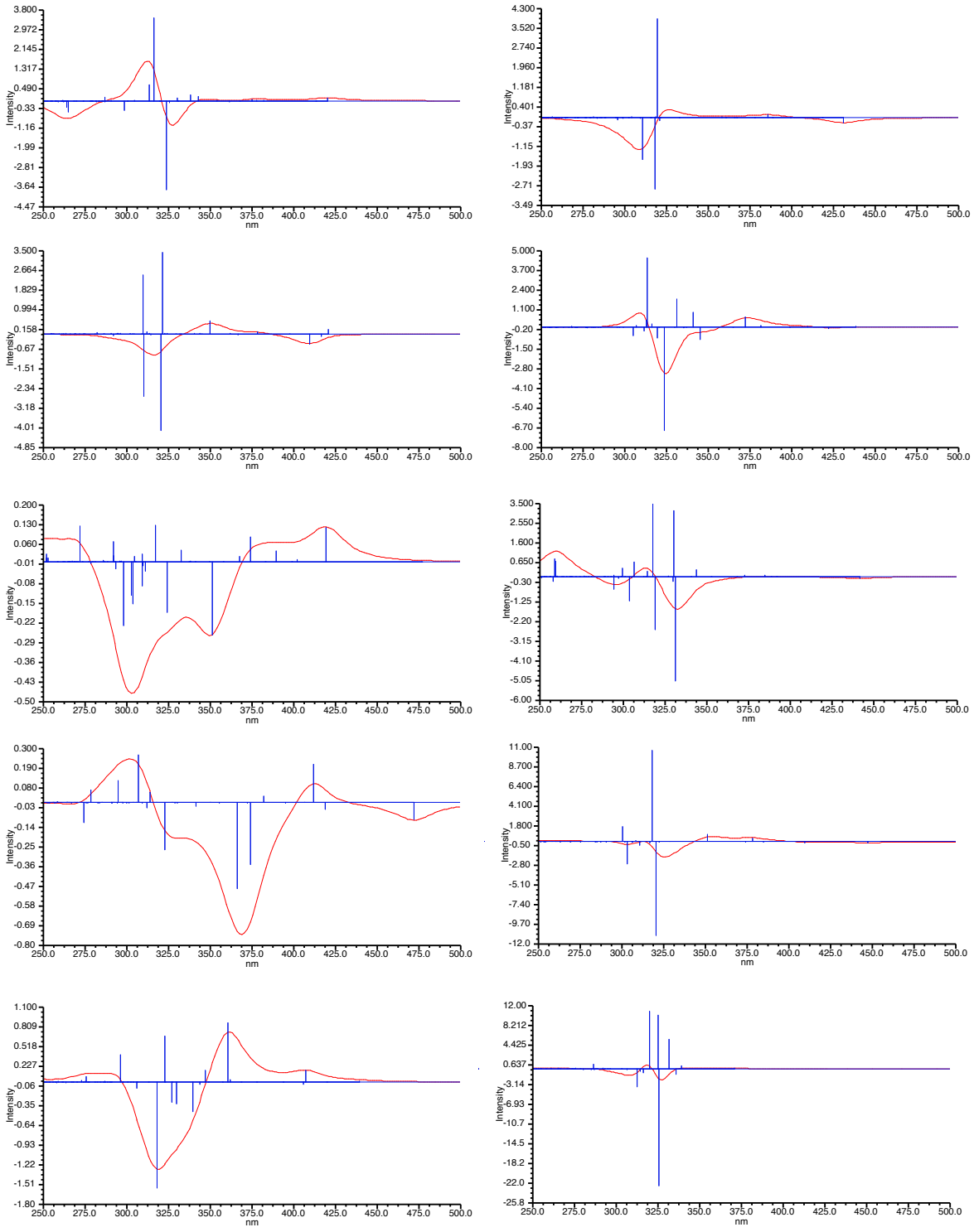


Figure 4.46. CD predictions of 4_1 knot using molecular dynamics *via* method b (described in Chapter 7).

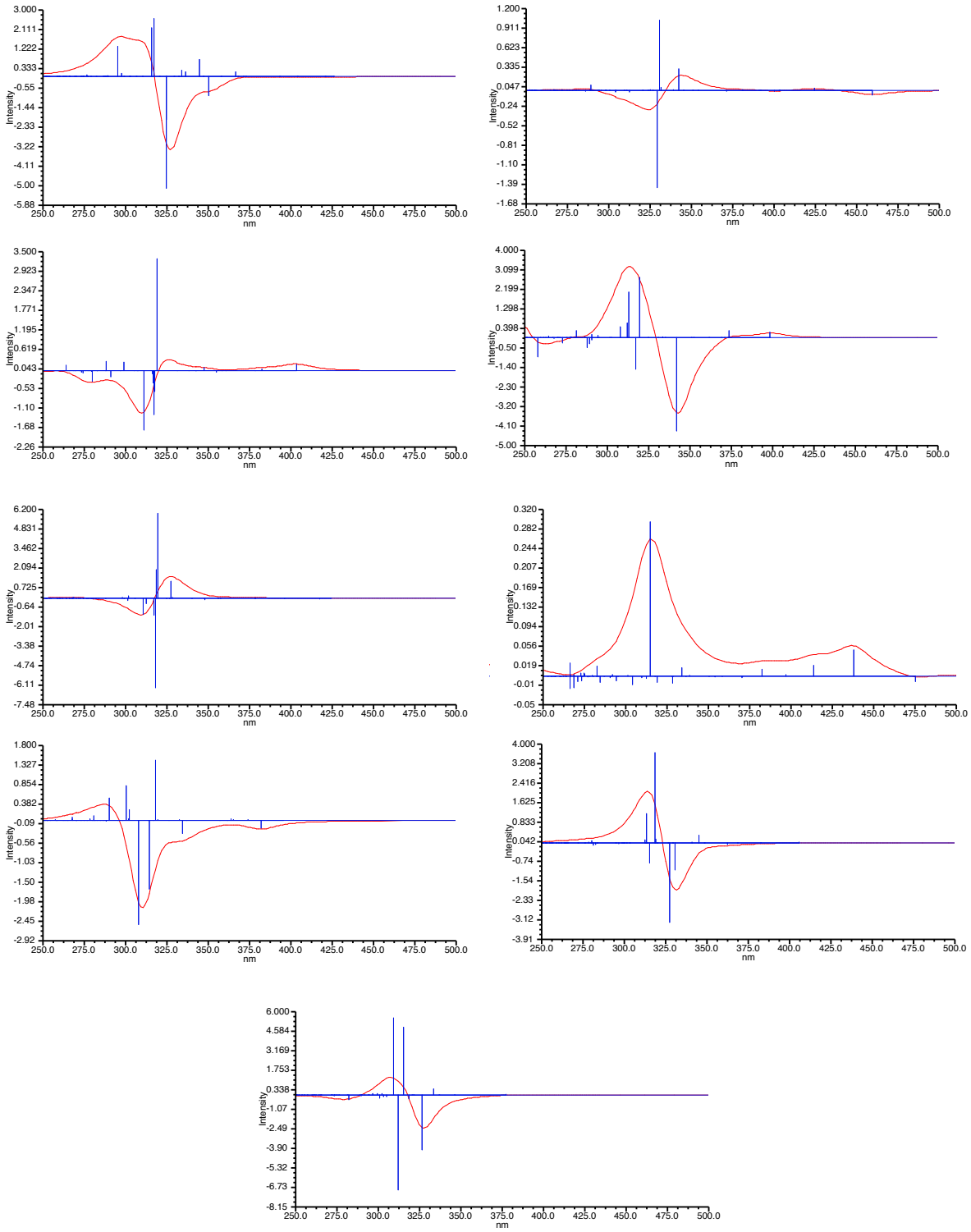
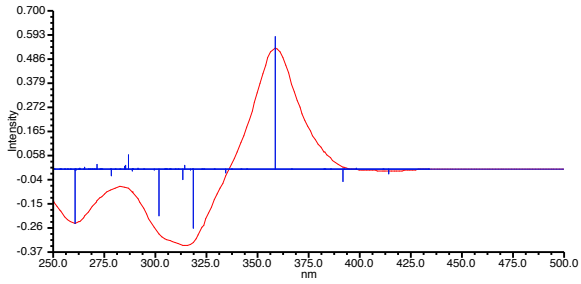
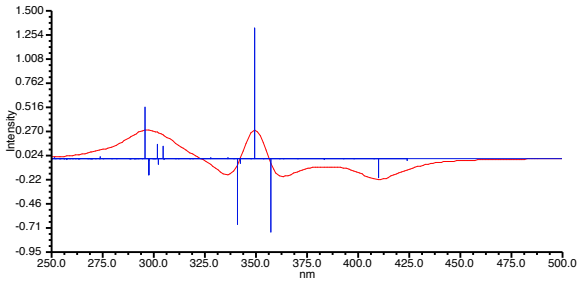
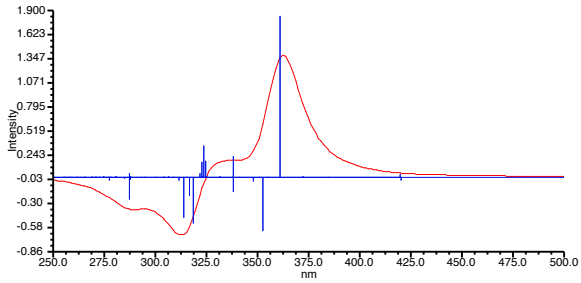
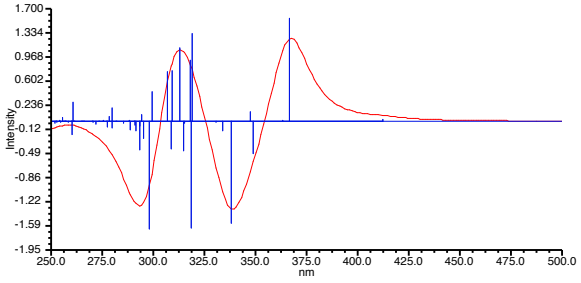


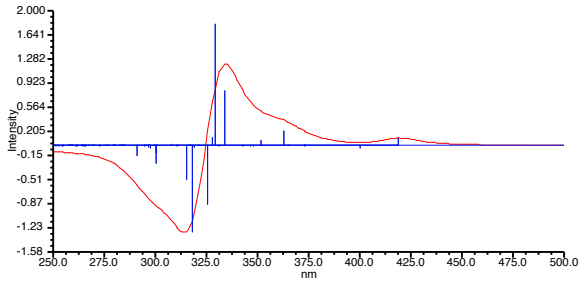
Figure 4.47. CD predictions of 4_1 knot using molecular dynamics *via* method c (described in Chapter 7).

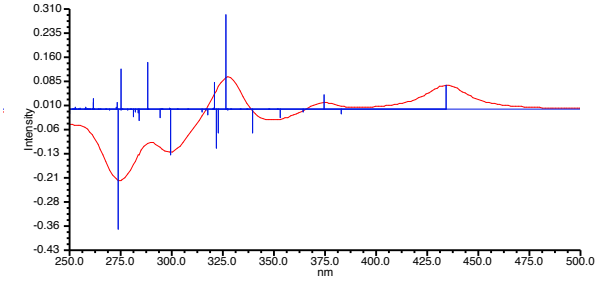

Intensity
nm

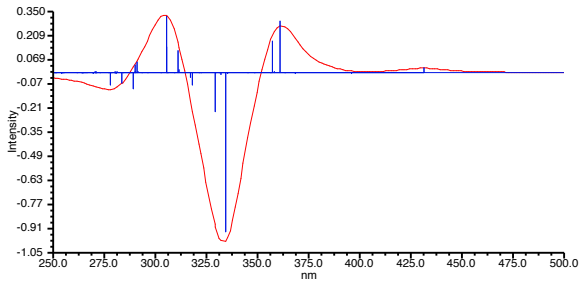

Intensity
nm

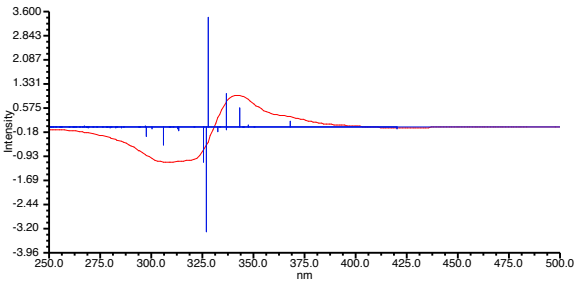

Intensity
nm

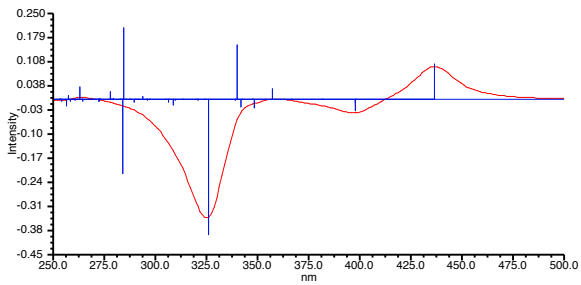

Intensity
nm

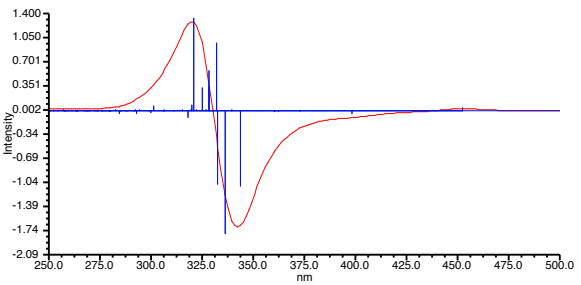

Intensity
nm

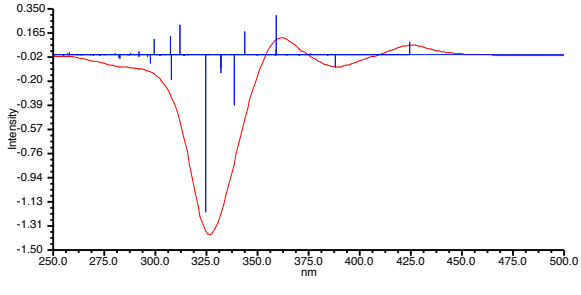

Intensity
nm

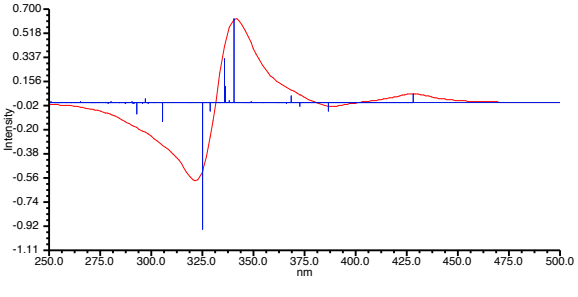

Intensity
nm


Intensity
nm


Intensity
nm


Intensity
nm


Intensity
nm


Intensity
nm

Intensity
nm

Intensity
nm

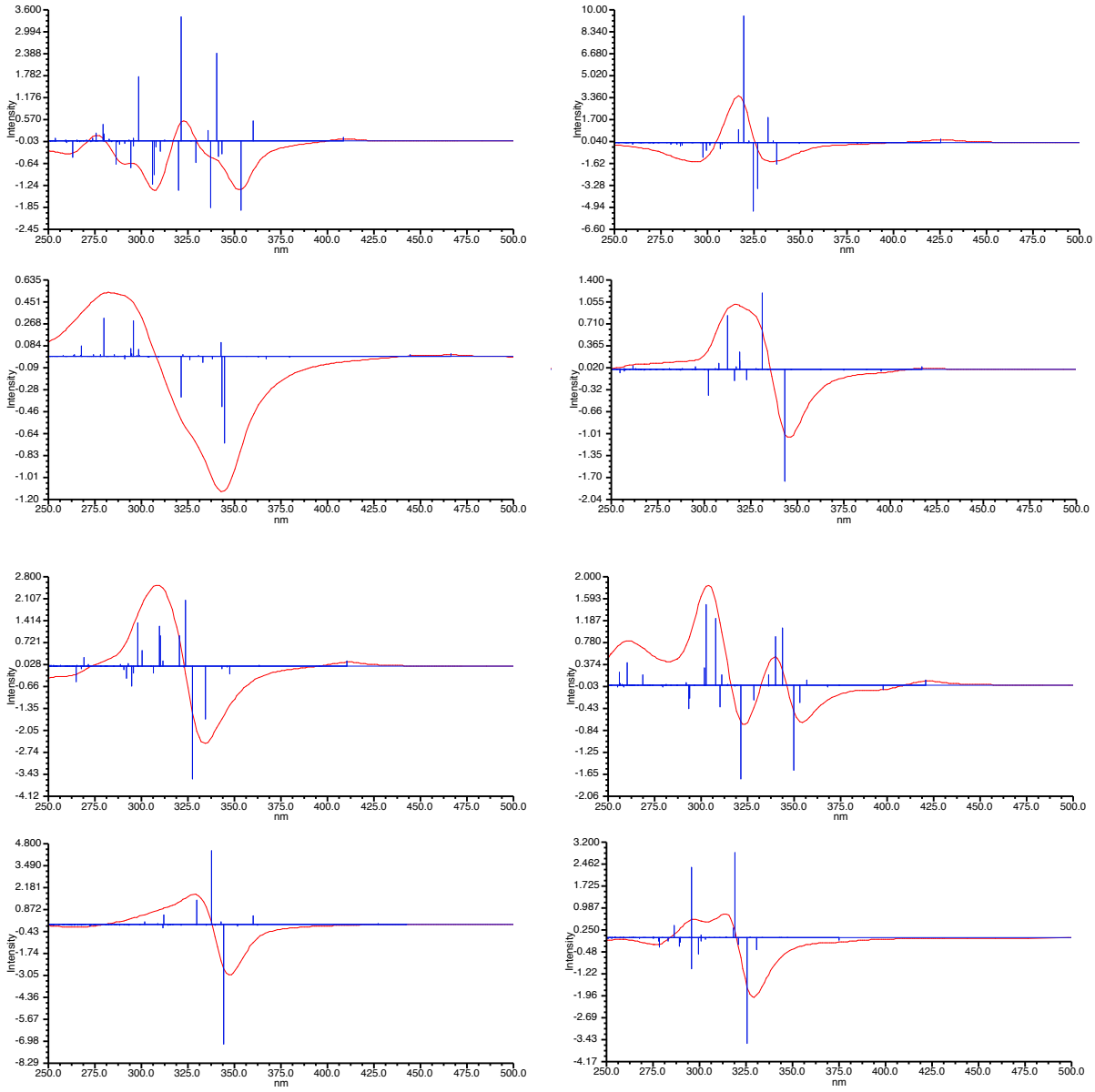


Figure 4.48. CD predictions of 5_1 knot using molecular dynamics *via* method a (described in Chapter 7).

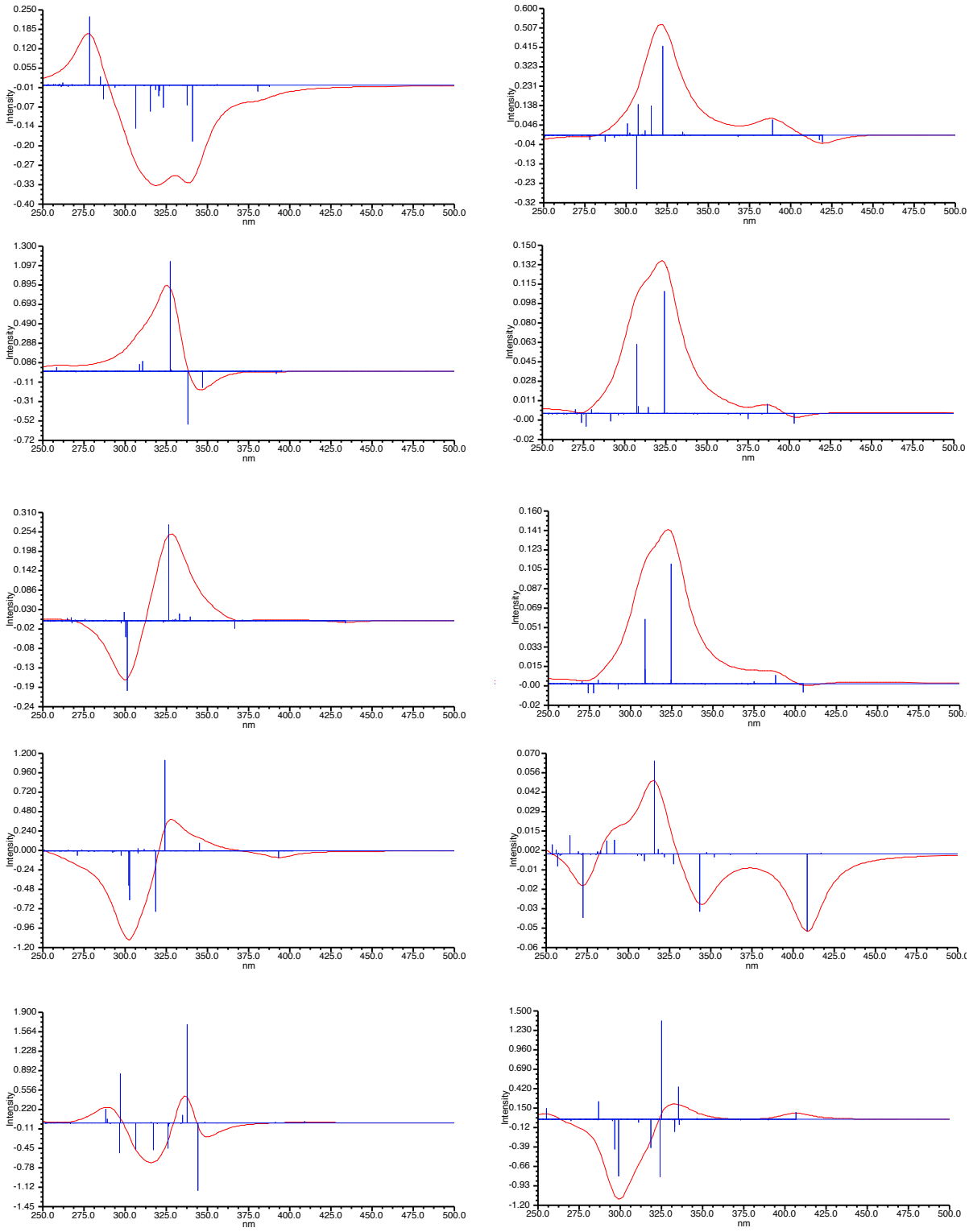


Figure 4.49. CD predictions of 5_1 knot using molecular dynamics *via* method b (described in Chapter 7).

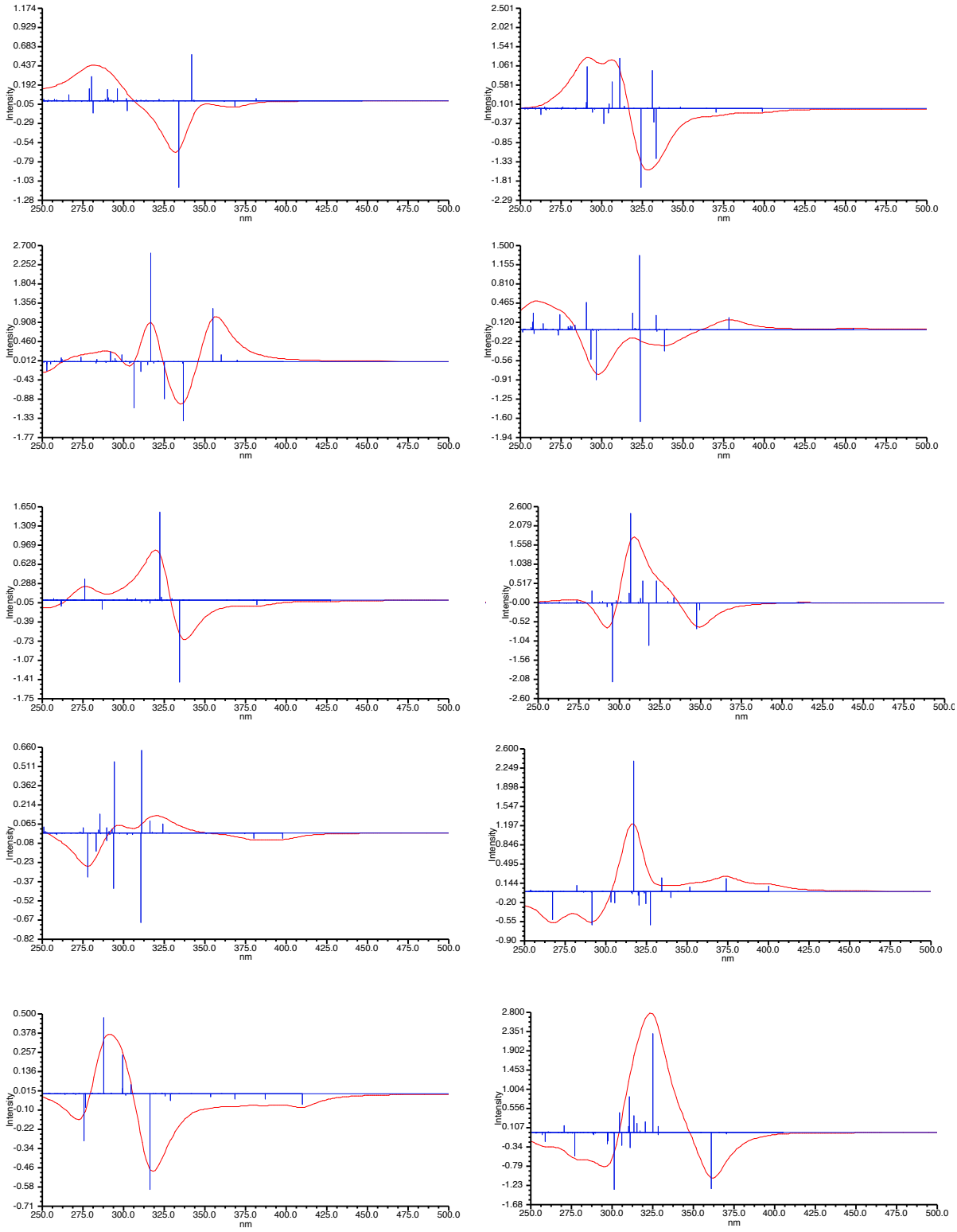
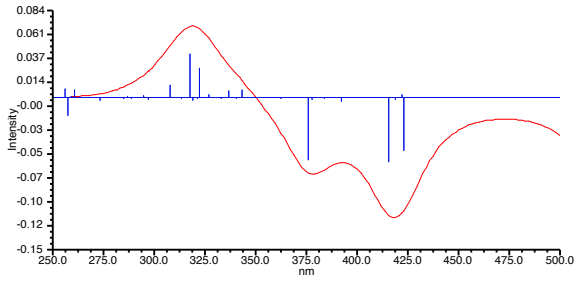
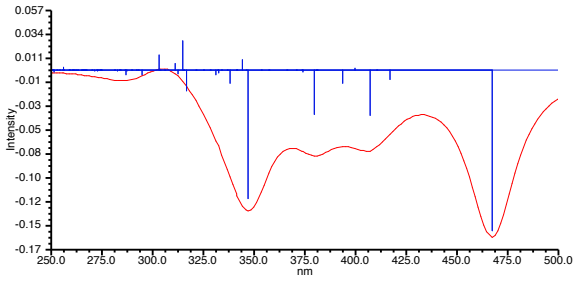
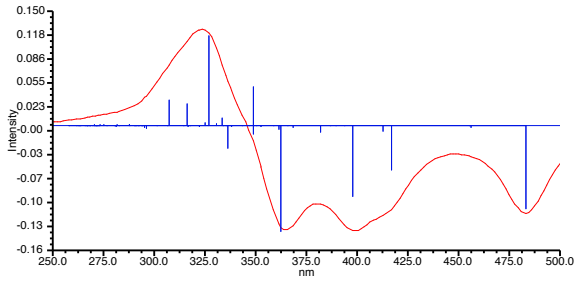
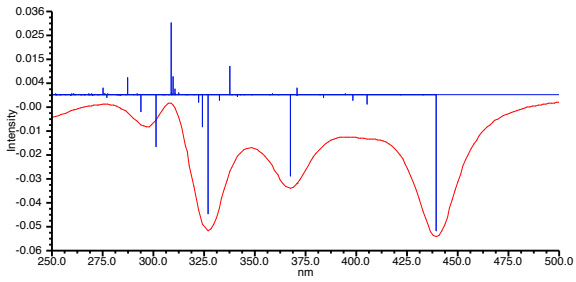


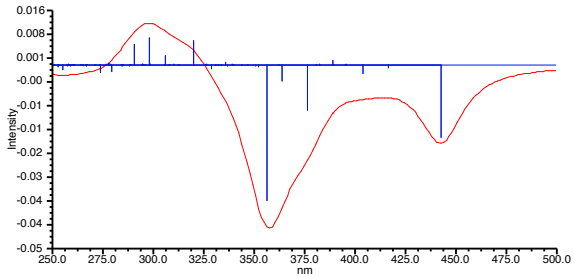
Figure 4.50. CD predictions of 5_1 knot using molecular dynamics *via* method c (described in Chapter 7).

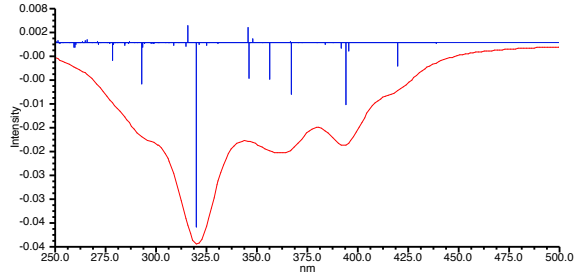

Intensity
nm

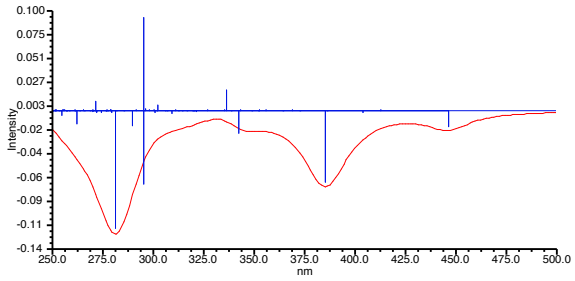

Intensity
nm

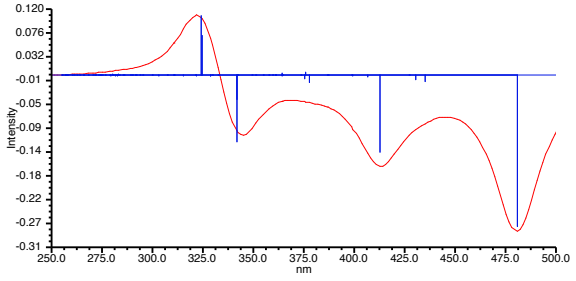

Intensity
nm

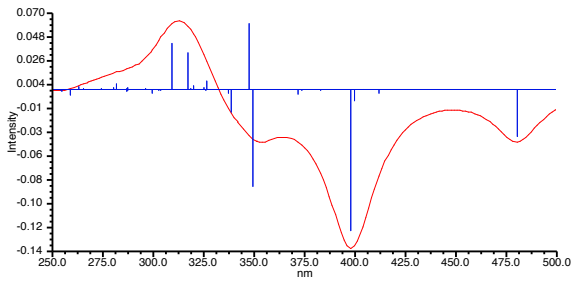

Intensity
nm

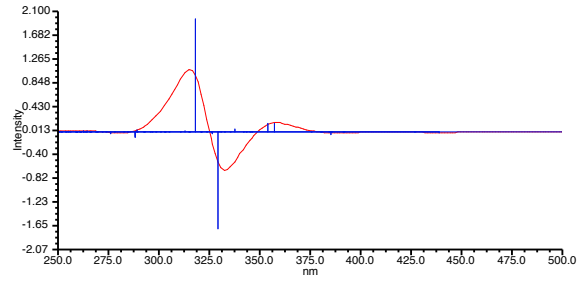

Intensity
nm

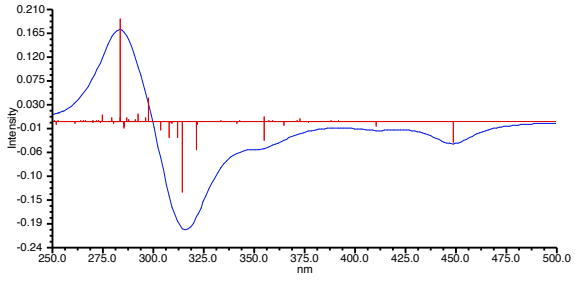

Intensity
nm

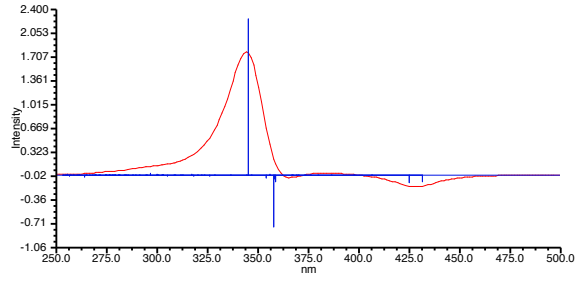

Intensity
nm


Intensity
nm


Intensity
nm


Intensity
nm


Intensity
nm


Intensity
nm

Intensity
nm

Intensity
nm

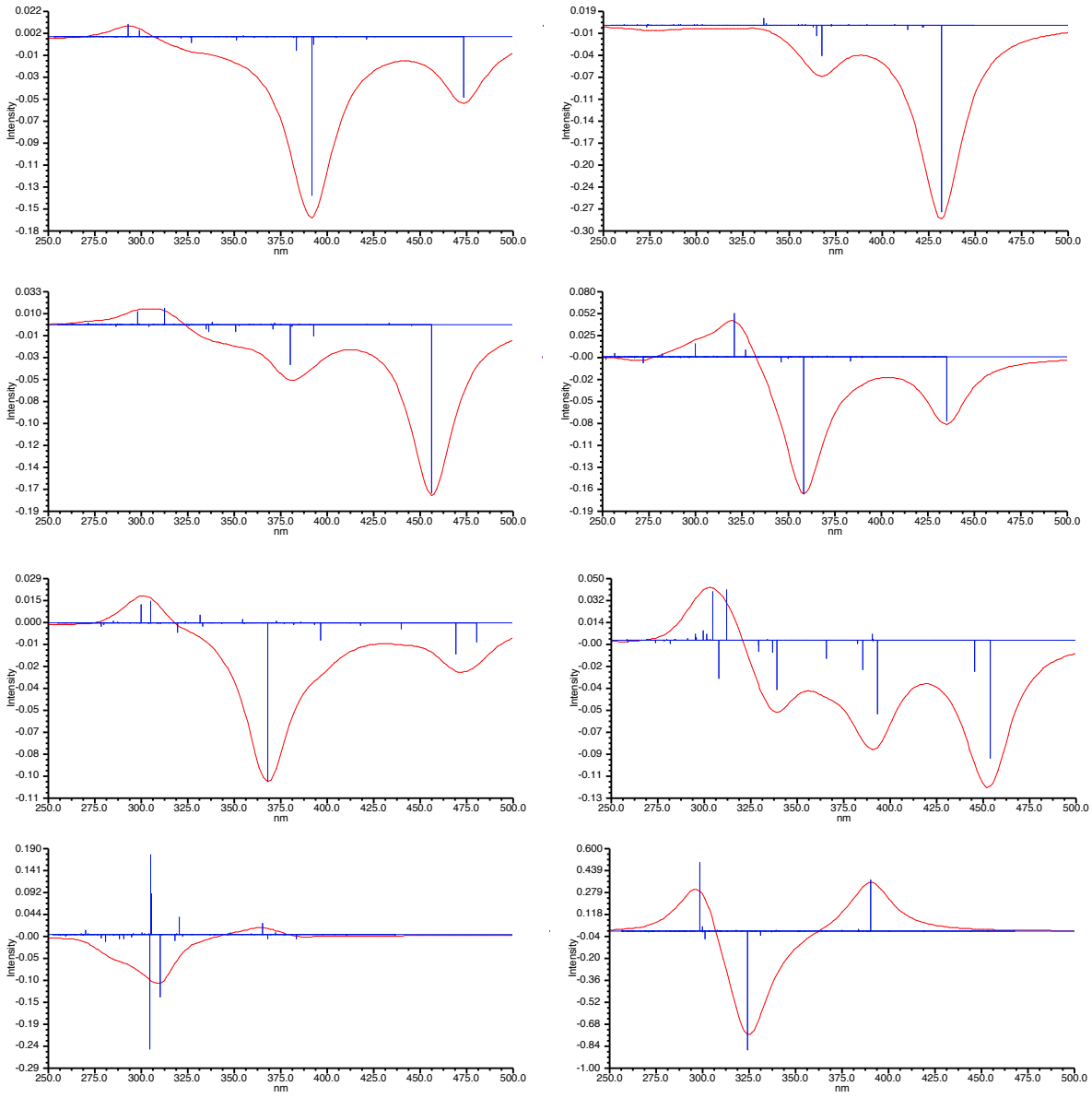


Figure 4.51. CD predictions of 5_2 knot using molecular dynamics *via* method a (described in Chapter 7).

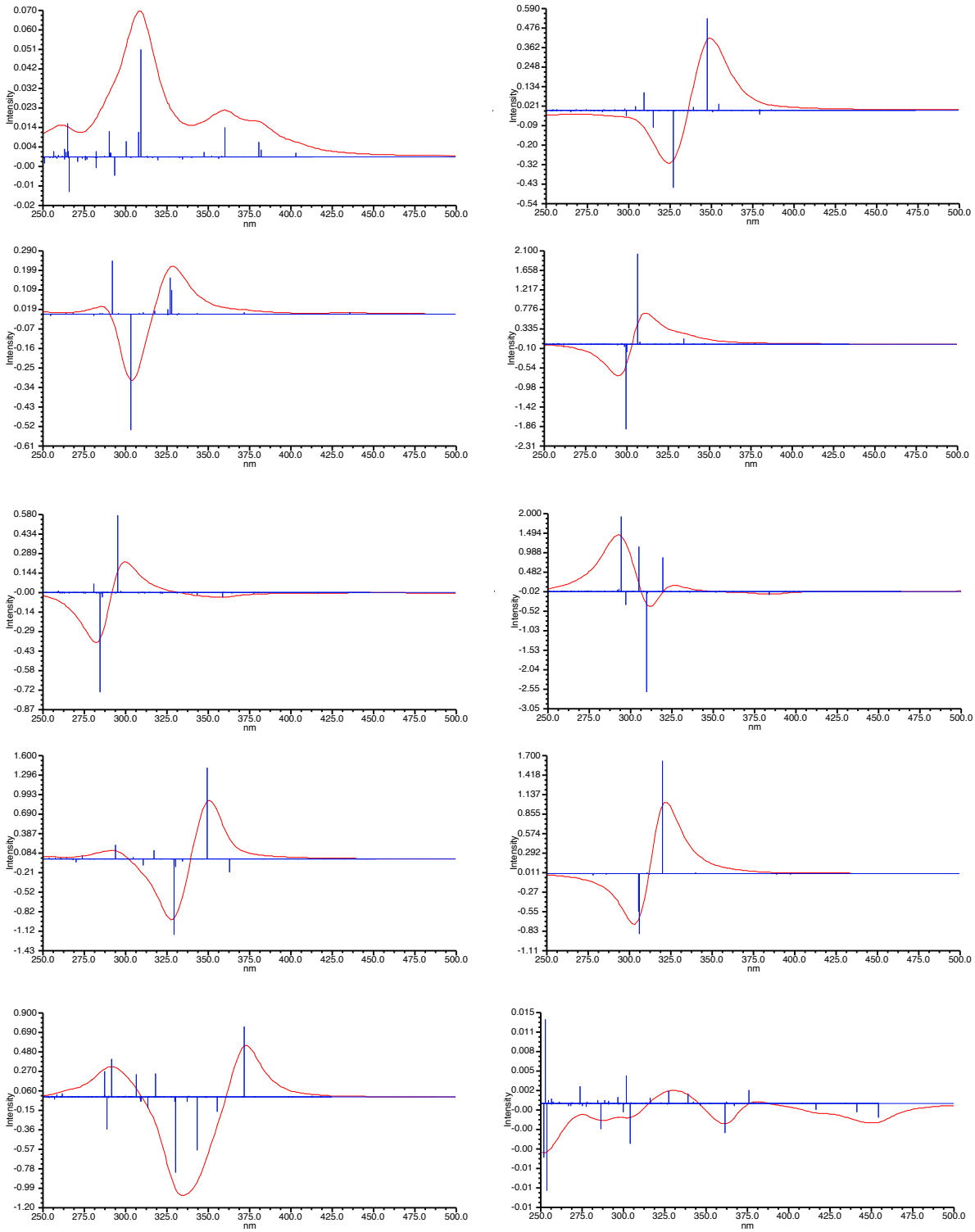


Figure 4.52. CD predictions of 5_2 knot using molecular dynamics *via* method b (described in Chapter 7).

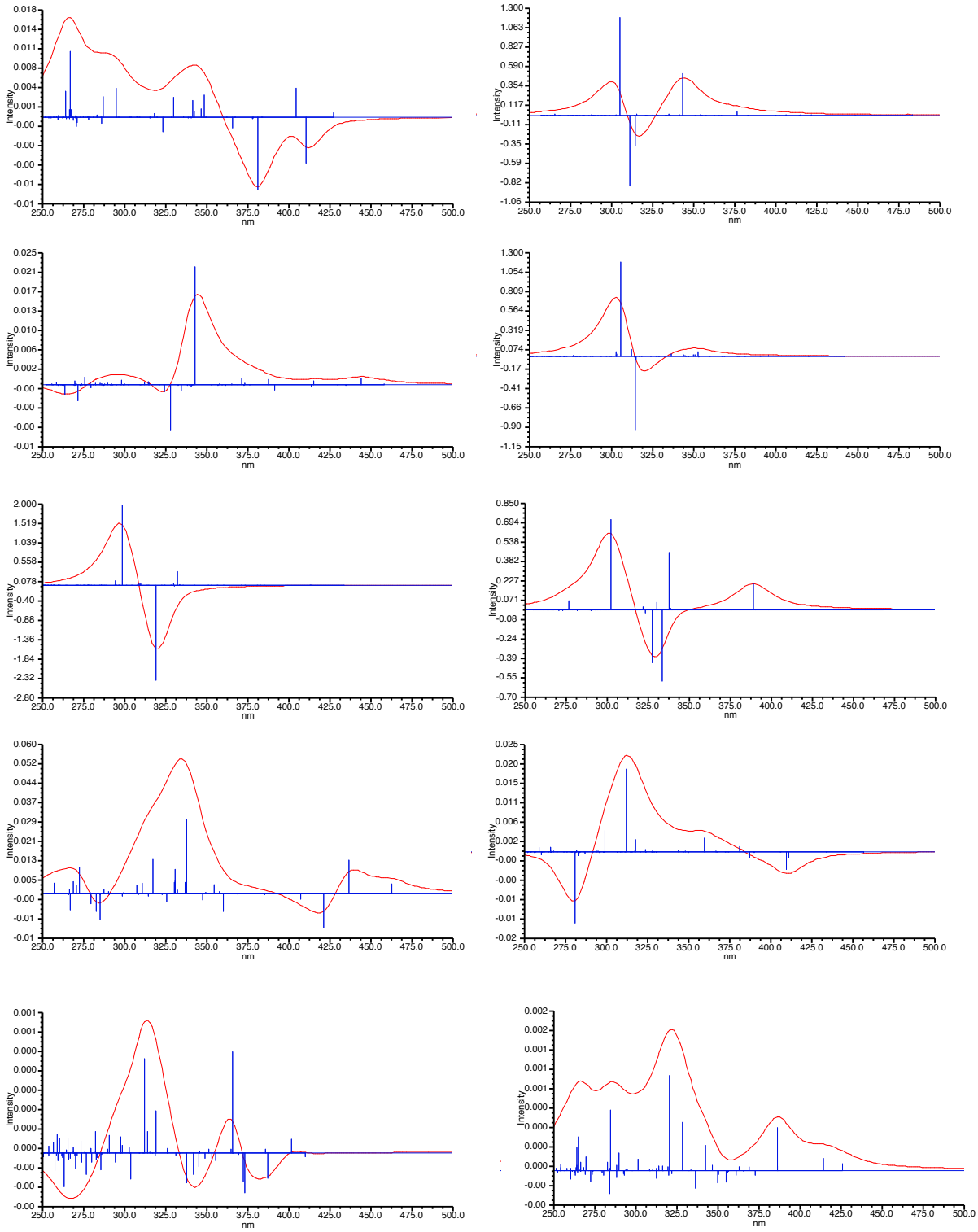


Figure 4.53. CD predictions of 5_2 knot using molecular dynamics *via* method c (described in Chapter 7).

Table 4.1. Closeness to fit value of some knotted conformations.

Structure	r	notes
4_1 knot	0.40	Has a C_2 symmetry
5_1 knot	0.45	Does not have a C_2 symmetry
5_2 knot	-0.02	Has a C_2 symmetry
3_1 knot	-0.03	Does not have a C_2 symmetry
4_1 knot- conf. 2	0.05	Does not have a C_2 symmetry
4_1 knot- conf. 3	-0.16	Does not have a C_2 symmetry
4_1 knot- conf. 4	0.11	Has a C_2 symmetry
5_2 knot conf. 2	-0.10	Has a C_2 symmetry

4.4. References

1 H. Y. Au-Yeung, G. D. Pantoş and J. K. M. Sanders, *Proc. Natl. Acad. Sci.*, 2009, **106**, 10466–10470.

2 H. Y. Au-Yeung, G. D. Pantoş and J. K. M. Sanders, *J. Am. Chem. Soc.*, 2009, **131**, 16030–16032.

3 H. Y. Au-Yeung, F. B. L. Cougnon, S. Otto, G. D. Pantoş and J. K. M. Sanders, *Chem. Sci.*, 2010, **1**, 567.

4 H. Y. Au-Yeung, G. D. Pantoş and J. K. M. Sanders, *Angew. Chem. Int. Ed.*, 2010, **49**, 5331–5334.

5 H. Y. Au-Yeung, G. D. Pantoş and J. K. M. Sanders, *J. Org. Chem.*, 2011, **76**, 1257–1268.

6 F. B. L. Cougnon, H. Y. Au-Yeung, G. D. Pantoş and J. K. M. Sanders, *J. Am. Chem. Soc.*, 2011, **133**, 3198–3207.

7 F. B. L. Cougnon, N. A. Jenkins, G. D. Pantoş and J. K. M. Sanders, *Angew. Chem. Int. Ed.*, 2012, **51**, 1443–1447.

8 N. Ponnuswamy, F. B. L. Cougnon, G. D. Pantoş and J. K. M. Sanders, *J. Am. Chem. Soc.*, 2014, **136**, 8243–8251.

9 A. R. Stefankiewicz, M. R. Sambrook and J. K. M. Sanders, *Chem. Sci.*, 2012, **3**, 2326.

10 A. R. Stefankiewicz and J. K. M. Sanders, *Chem. Commun.*, 2013, **49**, 5820.

11 N. Ponnuswamy, F. B. L. Cougnon, J. M. Clough, G. D. Pantoş and J. K. M. Sanders, *Science*, 2012, **338**, 783–785.

12 J. M. Clough, Cambridge, 2012.

13 H. Guenther and J. Guenther, *Chem. Rev.*, 1977, **77**, 599–637.

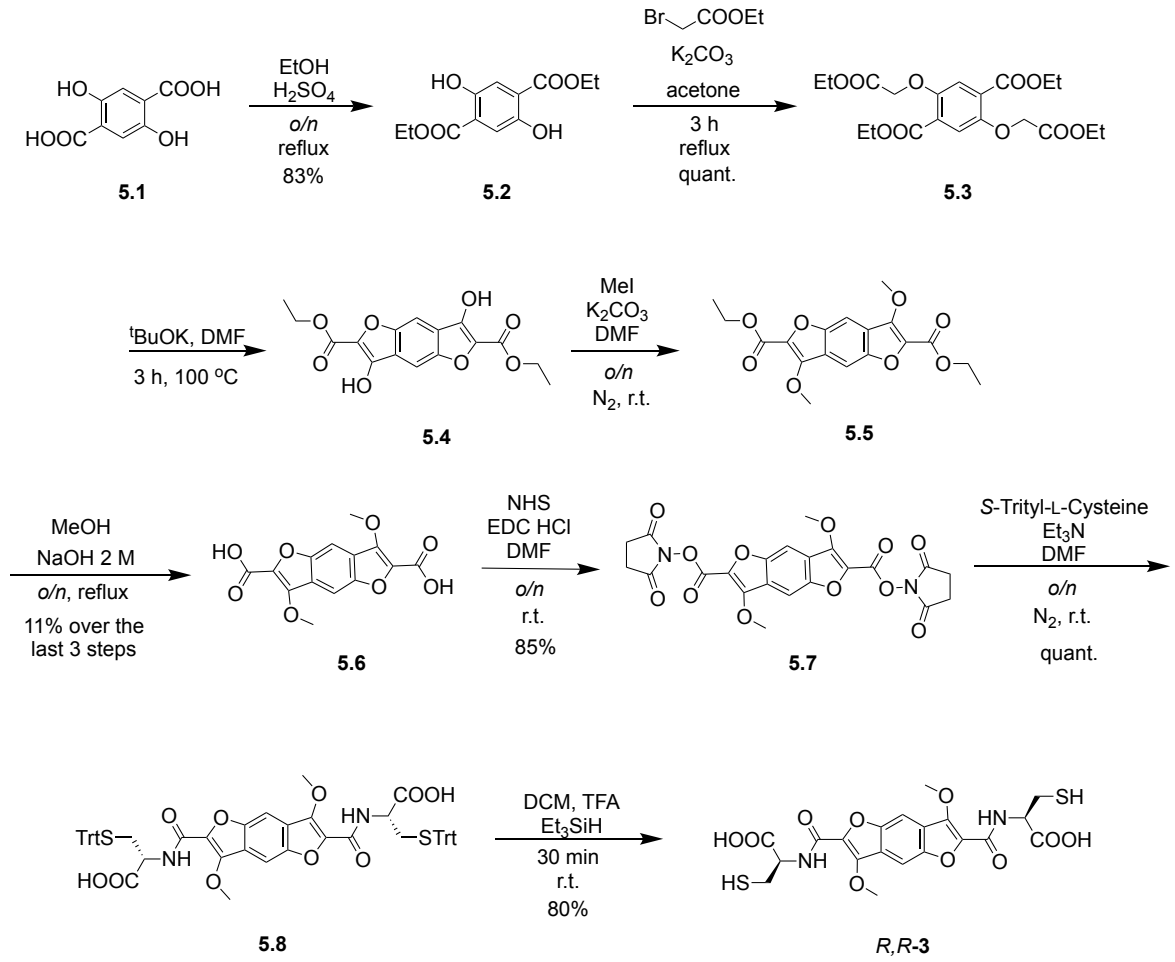
Chapter 5. Benzodifuran and Pyrroloindoles – new π-donor molecules

The remarkable results obtained with benzodithiophene (BDT) reported in Chapter 2 have led us to explore DCLs based on BDT analogues, in which the sulphur atom has been replaced by oxygen and nitrogen, respectively.

Benzodifuran (BDF) has been synthesised according to Scheme 5.1; this was the result of an original synthetic method development. After the esterification of the starting material (**5.1**) and the S_N2 reaction with ethyl bromoacetate **5.2**, the key step was the Claisen condensation, which afforded the benzodifuran ring **5.3**. Initially, the reaction was performed in THF, but it did not proceed as expected and the desired product did not form based on ^{1}H NMR analysis. **5.4** was obtained when the reaction solvent was changed to DMF because it is a better solvent for reactions involving polar transitions states. All the following steps have good to excellent yields, except for the synthesis of **5.5**, which has given a multitude of compounds leading to a mixture that could not be purified. The crude mixture of **5.5** was used for the synthesis of **5.6**, which was isolated in pure form. After NHS activation of **5.6**, the product was further reacted with enantiopure cysteine (*S*-Trityl-L-cysteine) to afford **5.8** – a benzodifuran derivative with cysteine appendages. The trityl deprotection has yielded the free thiol (*R,R*-**3**), which can be used as building block in DCLs. Many attempts to obtain single crystals of **5.8** were done without success. The computational data showed a Q_{zz} value of -21.1 B (thus having an electron-rich character), making it a good partner for the electron-deficient NDI.

5.1. Analysis of *R,R*-3 library

Firstly, DCLs of *R,R*-**3** on its own were set up in water at pH 8 without salt as well as in the presence of 1 M $NaNO_3$. The outcome of DCL containing *R,R*-**3** with no salt and in the presence of 1 M $NaNO_3$ is the homodimer (Figures 5.1 and 5.2). The compound was analysed using MS and MS/MS.



Scheme 5.1. The synthetic scheme of *R,R*-**3**.

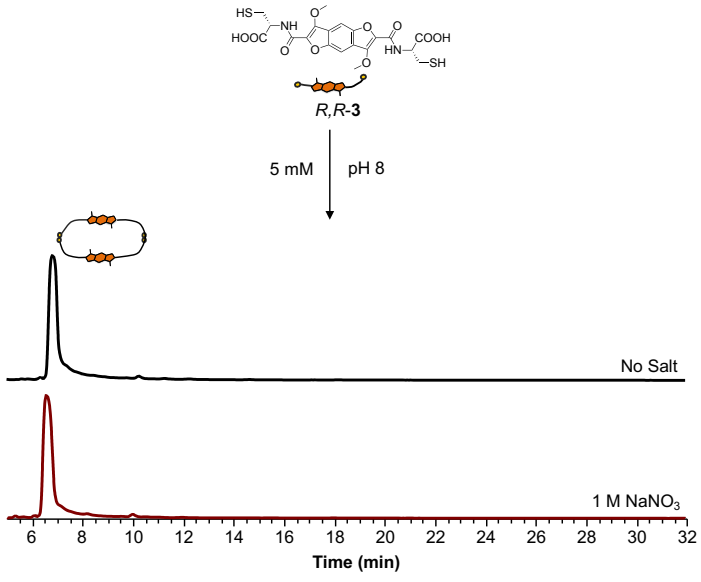


Figure 5.1. Reverse-phase HPLC analysis of *R,R*-**3** (5 mM total concentration) library without salt (top), and in the presence of 1 M $NaNO_3$ (bottom). Absorbances recorded at 375 nm. The unlabelled peaks did not ionise and could not be identified.

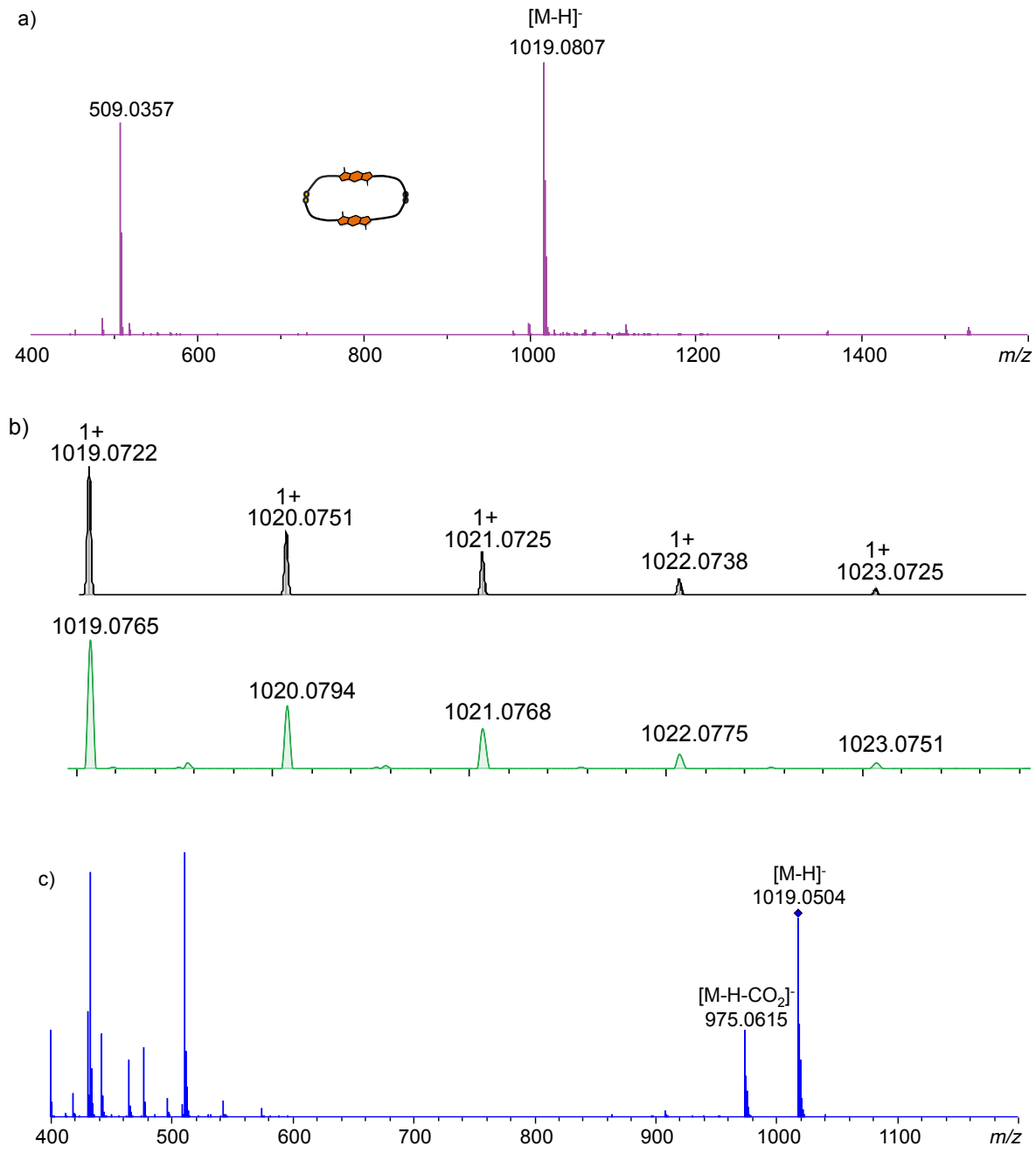


Figure 5.2. a) MS (*-ve*) of *R,R*-**3** dimer, b) expansion of region of *R,R*-**3** dimer and c) MS/MS (*-ve*) of *R,R*-**3** dimer.

5.2. Analysis of *R,R*-1 and *R,R*-3 libraries

A series of mixed DCLs with electron-deficient NDIs were set up. The first DCL was the homochiral one (*R,R*-**1** and *R,R*-**3**) in 1:1 ratio. The distribution of this DCL included the **Cat III *RRRR*** (a [2]catenane containing two *R,R*-**1** and two *R,R*-**3** units),

the homochiral homodimer (*R,R*-**3** dimer) and the homochiral heterodimer (Figure 5.3). The MS and MS/MS analyses of both compounds confirmed their structures. The MS/MS spectrum of the catenane showed no species with *m/z* higher than dimer, which was in line with the expected pattern (Figures 5.4 and 5.5).

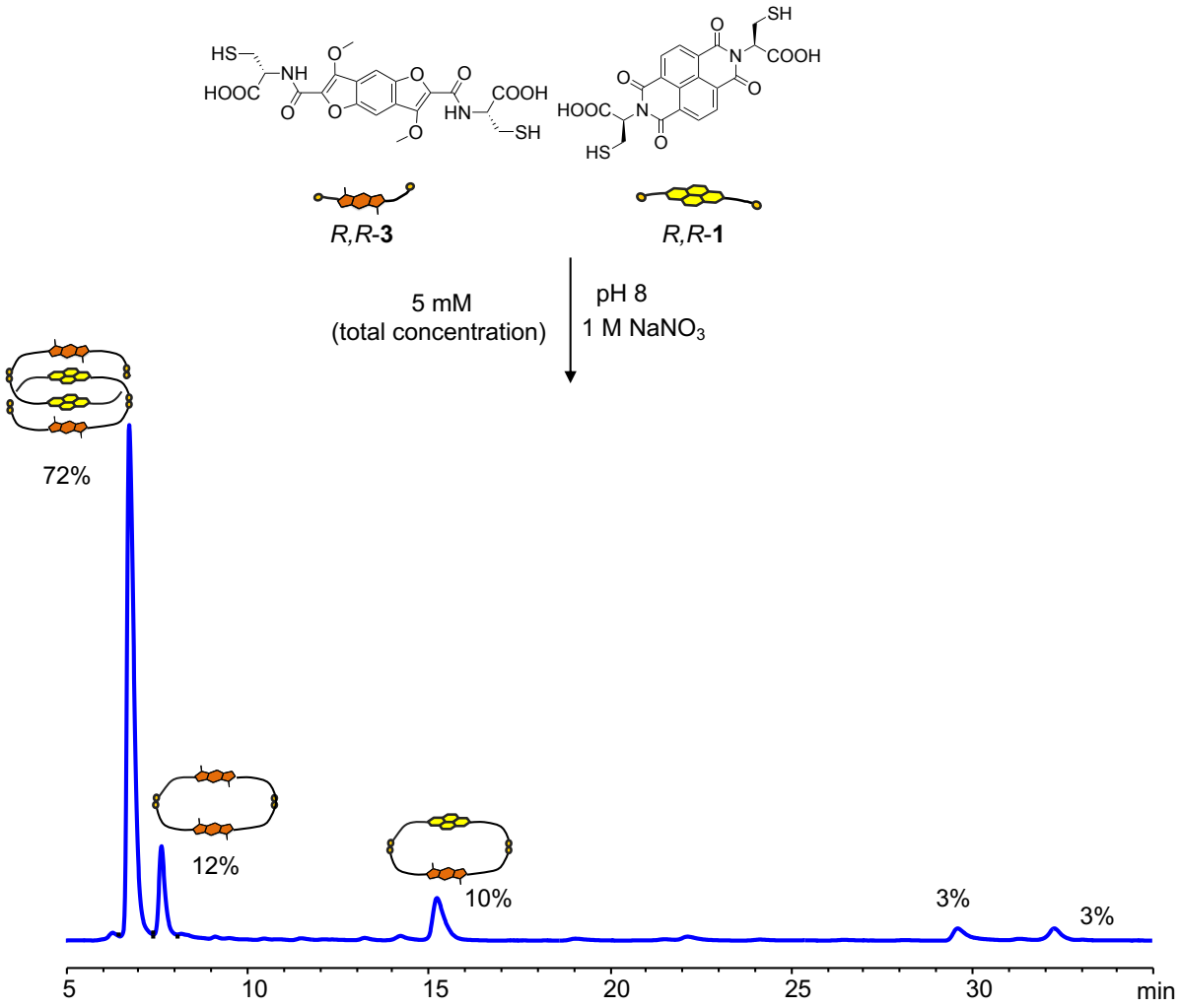


Figure 5.3. Reverse-phase HPLC analysis of *R,R*-**3**:*R,R*-**1** (1:1 molar ratio, 5 mM total concentration) library in the presence of 1 M $NaNO_3$. Absorbances recorded at 375 nm. The unlabelled peaks did not ionise and could not be identified.

5.2.1. Analysis of new *DAAD* [2]catenane

The [2]catenane was isolated using semi-preparative HPLC from a 3 mL library for further characterisation. The separation and the NMR analysis was performed by Hugo Wong (an exchange student in Pantoș' group from Hong Kong).

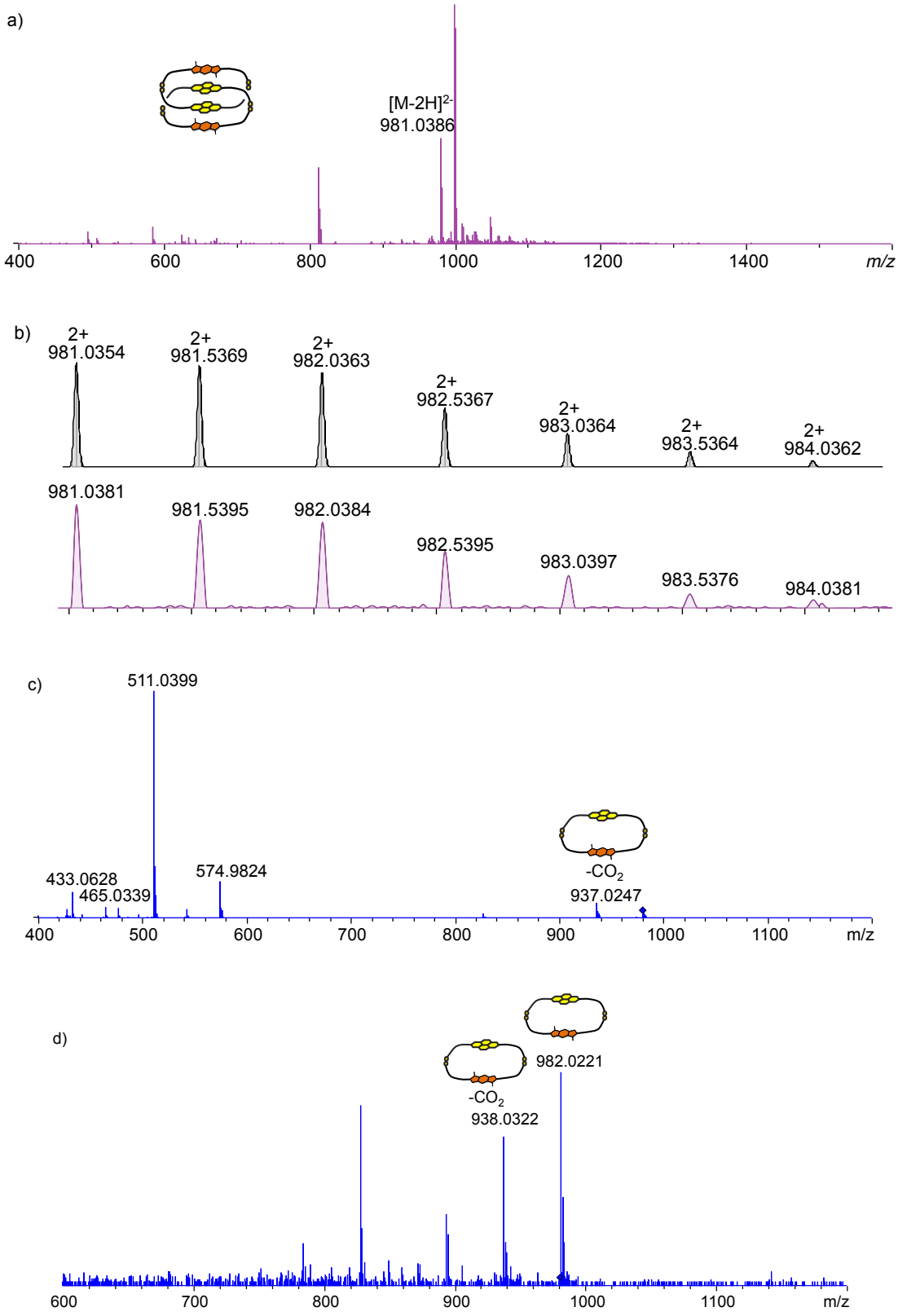


Figure 5.4. a) MS (-*ve*) of **Cat III *RRRR***, b) expansion of region of **Cat III *RRRR***, c) MS/MS (-*ve*) of of **Cat III *RRRR*** and d) expand of MS/MS (-*ve*) of of **Cat III *RRRR***.

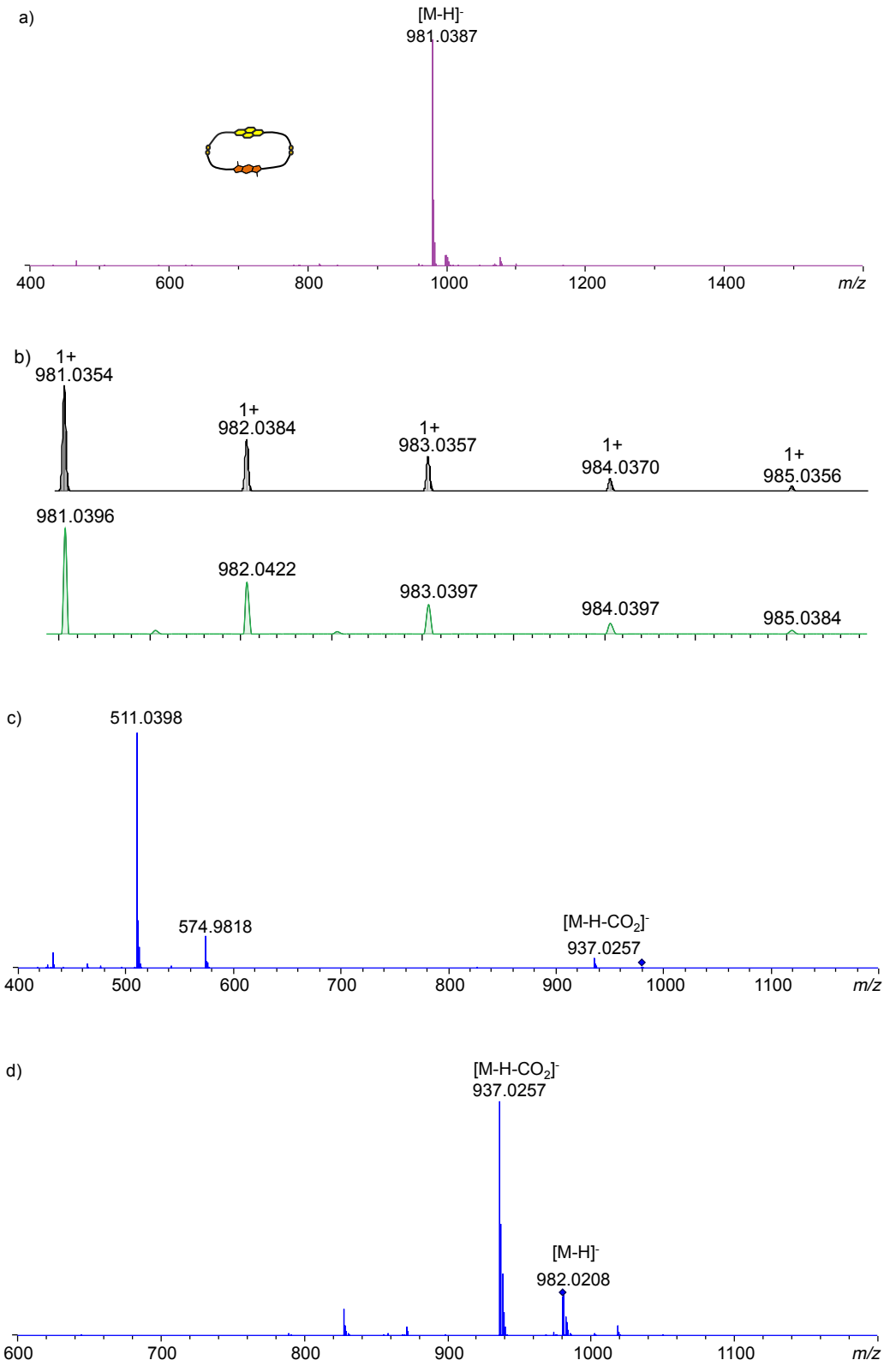


Figure 5.5. a) MS (*-ve*) of *R,R*-**3**, *R*,*R*-**1** heterodimer, b) expansion of region of *R,R*-**3**, *R*,*R*-**1** heterodimer, c) MS/MS (*-ve*) of *R,R*-**3**, *R*,*R*-**1** heterodimer and d) expansion of region of MS/MS (*-ve*) of *R,R*-**3**, *R*,*R*-**1** heterodimer.

The HPLC chromatogram of the library and that of the isolated compound are shown in Figure 5.6. The pattern corresponds to the aromatic protons on the desymetrised NDI core in ^{1}H NMR spectrum along with the upfield shifted BDF peak (Figure 5.7). The absorbance profile of the isolated catenane shows a combination of both chromophores (*R,R*-**1** and *R,R*-**3**) with broad peaks. The CD spectrum is characterised by a large Cotton effect (λ_{max} centred around 330 nm corresponding to DBF), indicating the rigid structure of **Cat III *RRRR***. This also shows an efficient chirality transfer from the chiral cysteine appendages onto the aromatic cores. VT UV-Vis and CD experiments suggest the decomposition of the molecule at high temperature (more than 65 ^{o}C), which can be due to photo and thermal degradation (Figure 5.8).

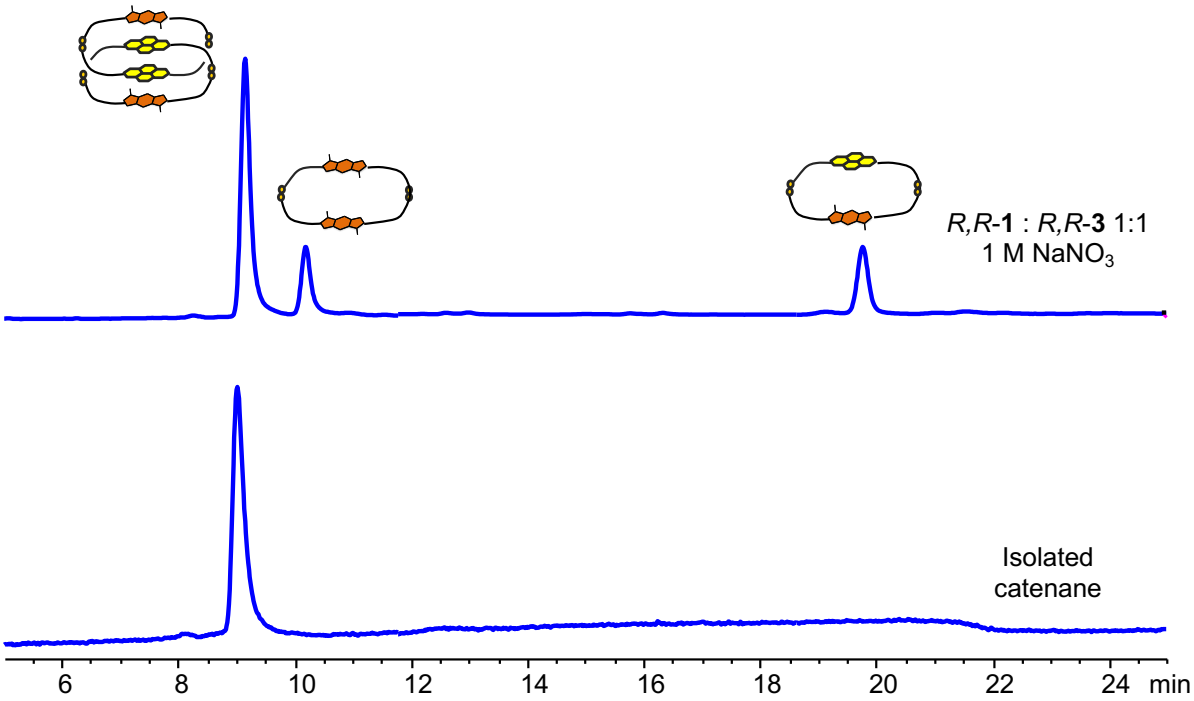


Figure 5.6. a) Reverse-phase HPLC analysis of *R,R*-**3**:*R,R*-**1** (1:1 molar ratio, 5 mM total concentration) library in the presence of 1 M $NaNO_3$ (top trace) and b) Reverse-phase HPLC analysis of **Cat III *RRRR*** (bottom trace). Absorbances recorded at 375 nm.

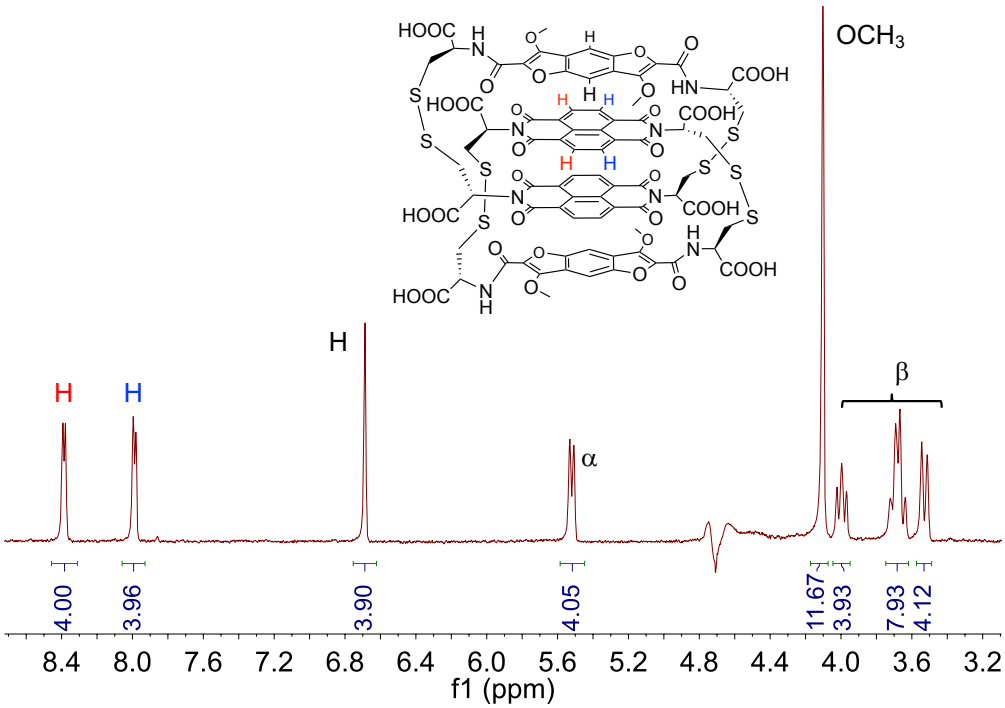


Figure 5.7. ^{1}H NMR spectrum (500 MHz) of **Cat III *RRRR***. The residual solvent (H_2O) was referenced at 4.79 ppm.

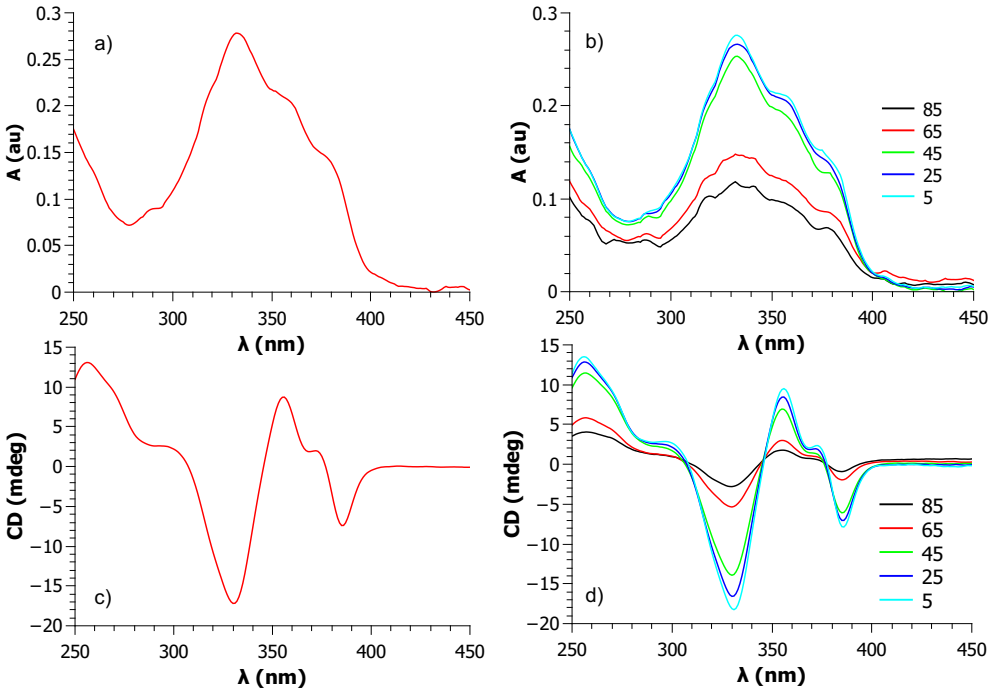


Figure 5.8. a) UV-Vis spectrum of **Cat III *RRRR*** at 23 °C (296 K), b) VT UV-Vis spectra of **Cat III *RRRR*** between 5 – 85 °C at the specified temperatures, c) CD spectrum of **Cat III *RRRR*** at 23 °C (296 K) and d) VT CD spectra of **Cat III *RRRR***, between 5 – 85 °C at the specified temperatures. The solvent was a mixture of water-acetonitrile with 0.1% TFA.

5.3. Analysis of *S,S*-1 and *R,R*-3 libraries

The next DCL analysed was the heterochiral DCL (*S,S*-**1** and *R,R*-**3**). The distribution of the library at equilibrium shows the formation of another [2]catenane (**Cat IV *RSRR***), and dimers (the homochiral homodimers, *R,R*-**3** dimer, and the heterochiral heterodimer – Figure 5.9). The MS and MS/MS analyses confirm the catenated structure for **Cat IV *RSRR***, and the nature of the other dimers (Figures 5.10 and 5.11). The analysis also shows that the catenane is composed of one *S,S*-**1** and three *R,R*-**3** units.

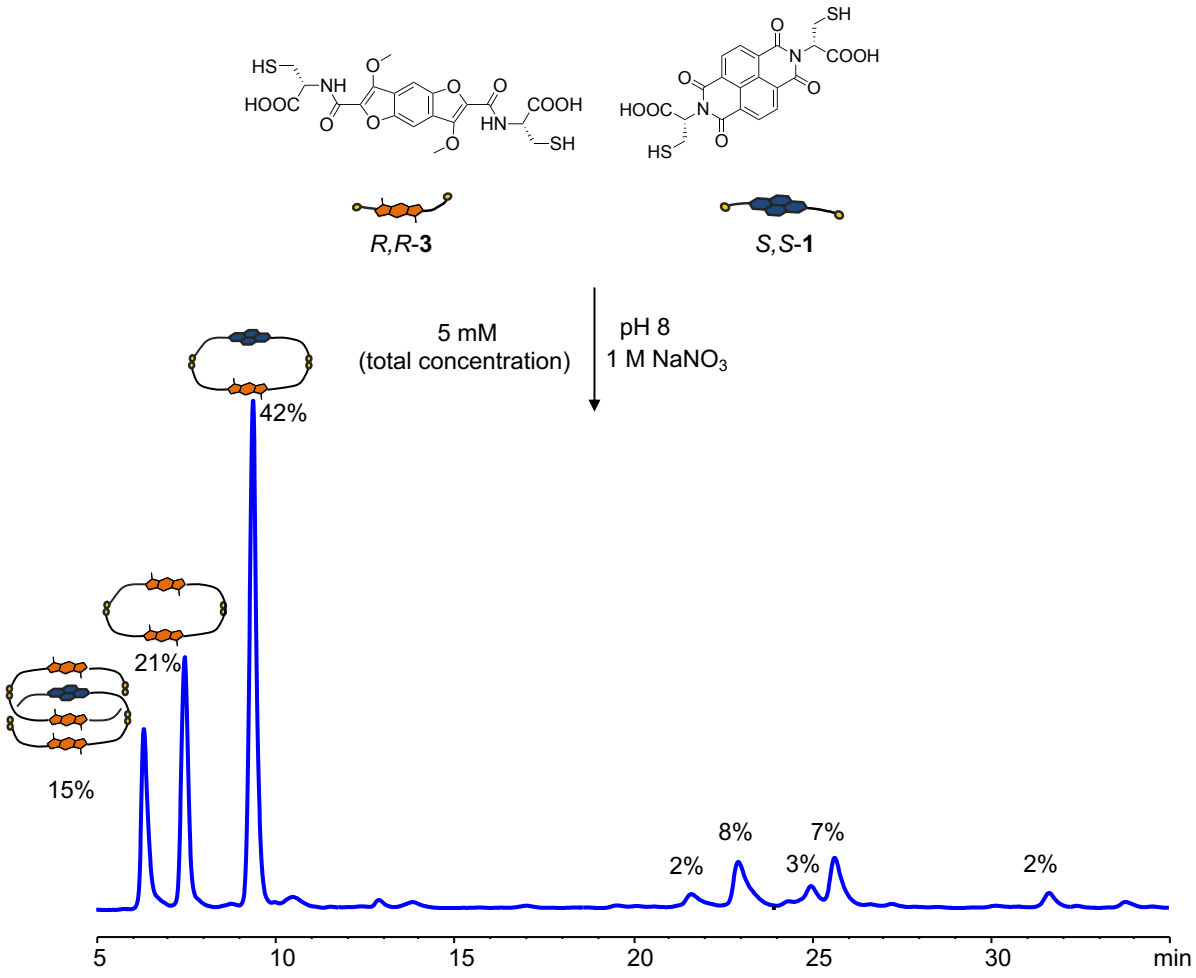


Figure 5.9. Reverse-phase HPLC analysis of *R,R*-**3**:*S,S*-**1** (1:1 molar ratio, 5 mM total concentration) library in the presence of 1 M $NaNO_3$. Absorbances recorded at 375 nm. The unlabelled peaks did not ionise and could not be identified.

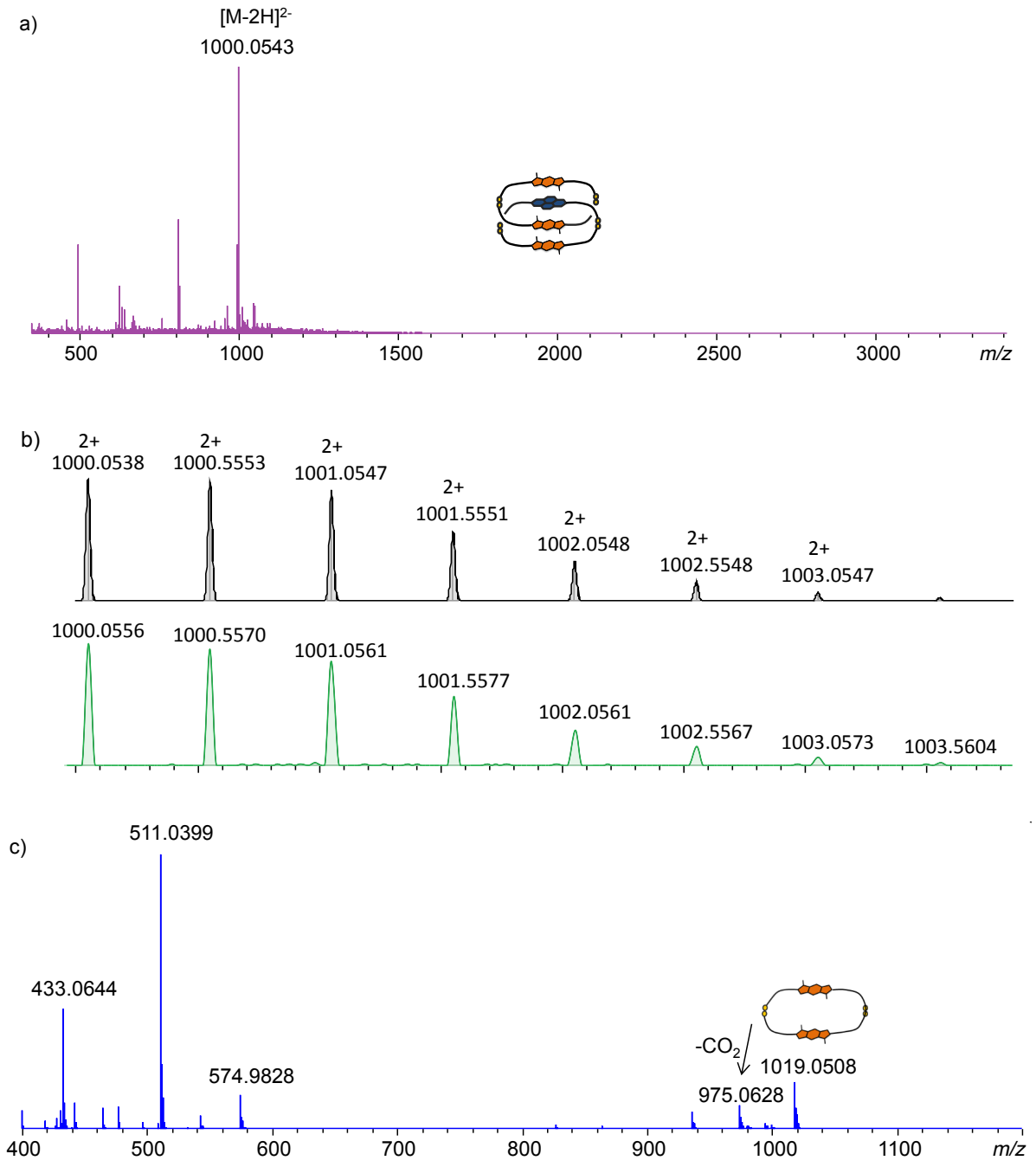


Figure 5.10. a) MS (-*ve*) of **Cat IV *RSRR***, b) expansion of region of **Cat IV *RSRR*** and c) MS/MS (-*ve*) of **Cat IV *RSRR***.

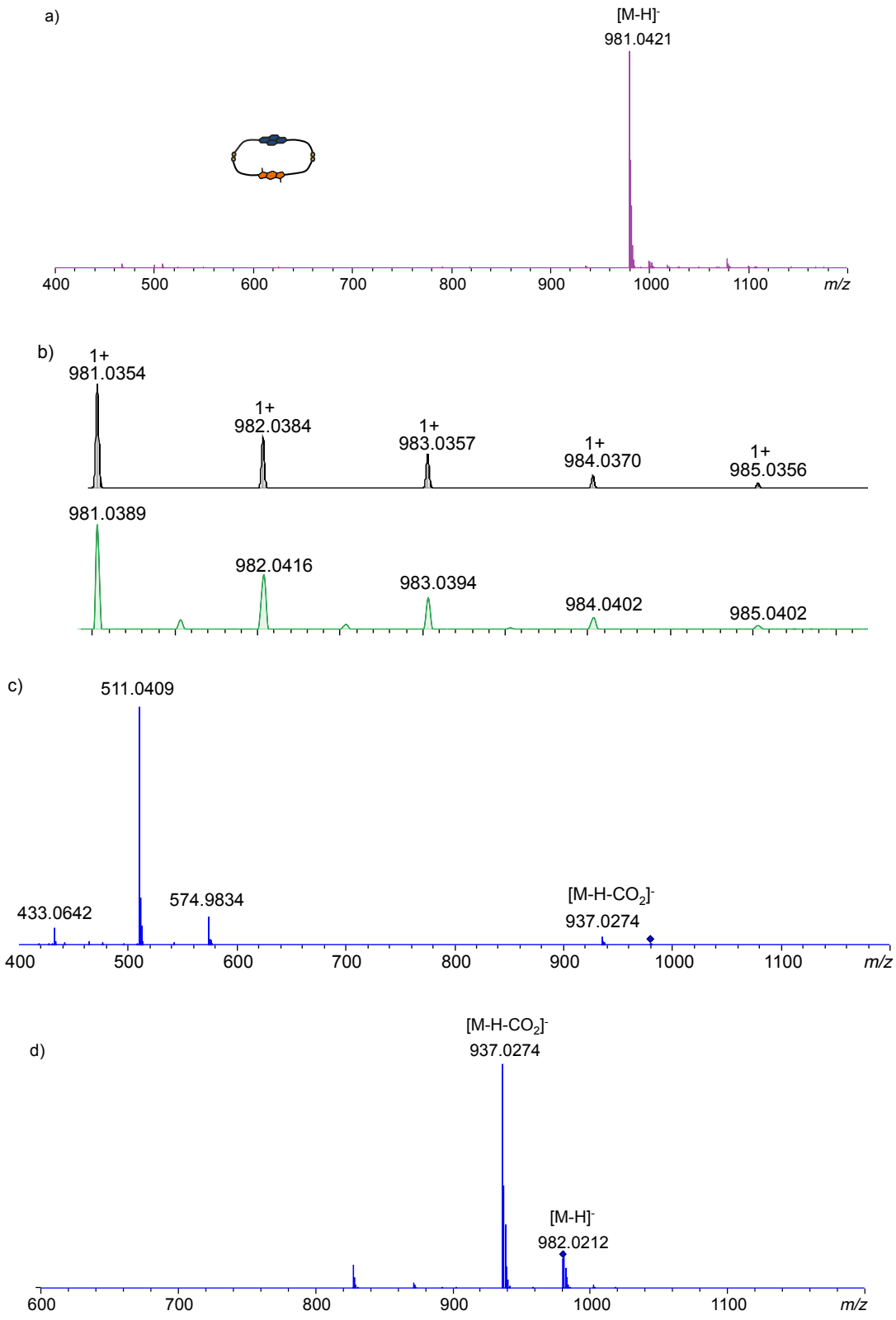


Figure 5.11. a) MS (*-ve*) of *R,R*-**3**, *S,S*-**1** heterochiral heterodimer, b) expansion of region of *R,R*-**3**, *S,S*-**1** heterochiral heterodimer, c) MS/MS (*-ve*) of *R,R*-**3**, *S,S*-**1** heterochiral heterodimer and d) expansion of MS/MS (*-ve*) of *R,R*-**3**, *S,S*-**1** heterochiral heterodimer.

5.3.1. Analysis of new *DAAD* [2]catenane

This [2]catenane was also isolated using semi-preparative HPLC from a 4 mL library. The separation and the NMR analysis was performed by Hugo Wong. The HPLC chromatograms of the library and that of isolated compound are provided in Figure 5.12. The analysis of **Cat IV *RSRR*** by ^{1}H NMR spectroscopy (Figure 5.13) outlines a complicated pattern consistent with the presence of one *S,S*-**1** and *R,R*-**3** moieties. There are the two doublets that come from the aromatic protons of the desymmetrised NDI cored and three singlets in the aromatic region. They correspond of the BDF core, the most shielded singlet being attributed to the inner BDF core. The UV-Vis spectrum of this [2]catenane shows the typical absorption profile observed for the BDF so far (λ_{max} around 330 nm), while the CD profile displays a large bisignate optical response. This suggests an effective chirality transfer from cysteine onto the BDF core, and so, the rigid structure of **Cat IV *RSRR***. CD and UV-Vis VT experiments suggest no decomposition of this catenane. This is in contrast with **Cat III *RRRR***, which decomposed (Figure 5.14) under an identical experimental setup. This highlights the difference in stabilities of *DADA vs. DADD* stacking, indicating that a stack containing more *R,R*-**3** units leads to a stable structure at photo- and thermal degradation. An example from literature shows a derivative of benzodifuran has a decomposition temperature more than 300 ^{o}C, which is in accordance with our supposition.[1]

The last part of the discussion focuses on the behaviour of the racemic mixture of NDI (*R,R*-**1** and *S,S*-**1**) with *R,R*-**3**. Initially, the library distribution shows the formation of catenanes **Cat III *RRRR*** and **Cat IV *RSRR*** along with the already mentioned three dimers (*R,R*-**3** dimer, *R,R*-**1**:*R,R*-**3** dimer and *S,S*-**1**:*R,R*-**3** dimer). The chromatogram also shows two new peaks, which correspond to two tetramers (based on MS and MS/MS analyses) – Figures 5.15 – 5.17. It is essential to notice that in this case there are no erosions in chirality (dr); only **Cat III *RRRR*** and **Cat IV *RSRR*** were formed.

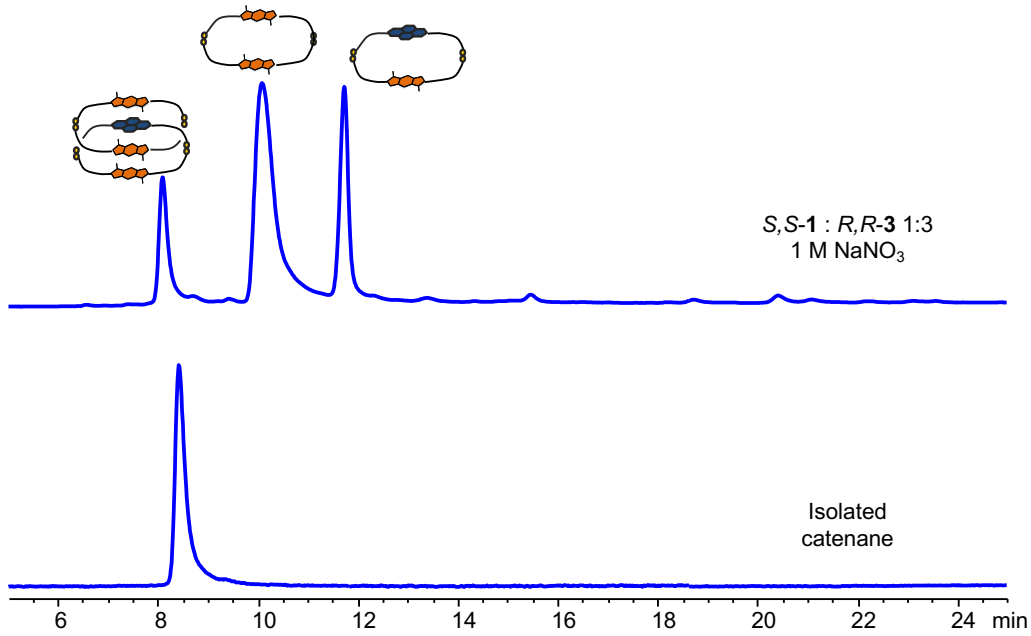


Figure 5.12. a) Reverse-phase HPLC analysis of *R,R*-**3**:*R,R*-**1** (3:1 molar ratio, 5 mM total concentration) library in the presence of 1 M $NaNO_3$ (top trace) and b) Reverse-phase HPLC analysis of **Cat IV *RSRR*** (bottom trace). Absorbances recorded at 375 nm.

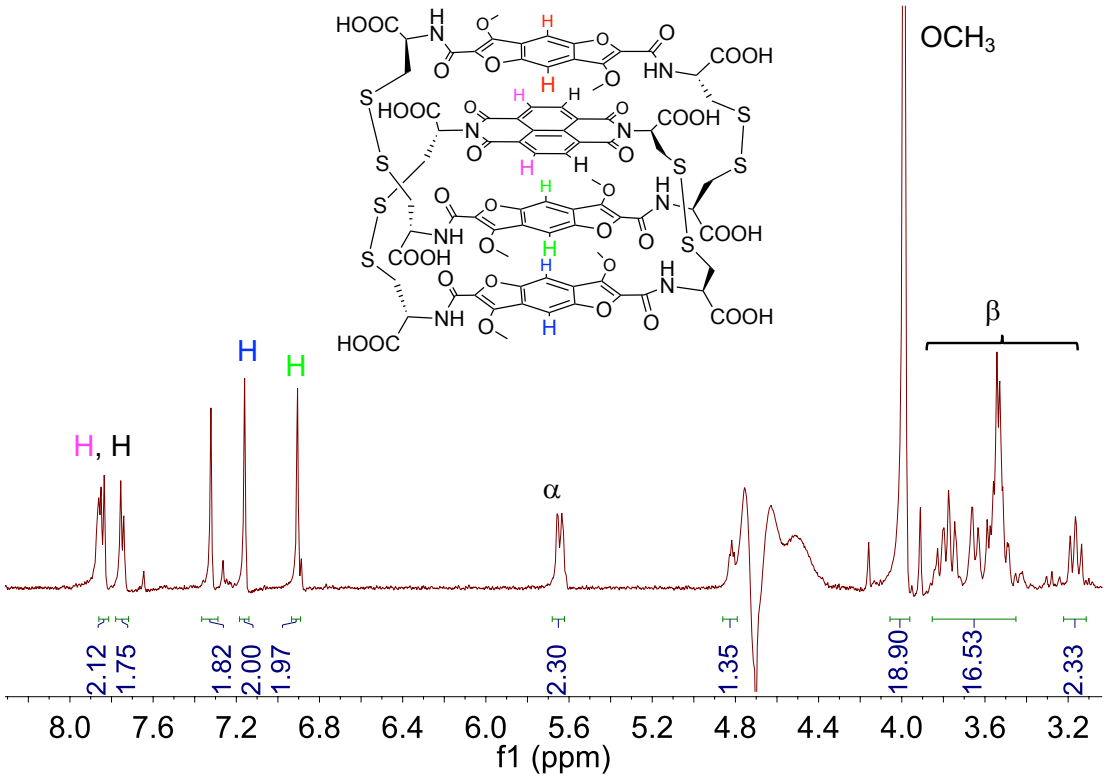


Figure 5.13. ^{1}H NMR spectrum (500 MHz) of **Cat IV *RSRR***. The residual solvent (H_2O) was referenced at 4.79 ppm.

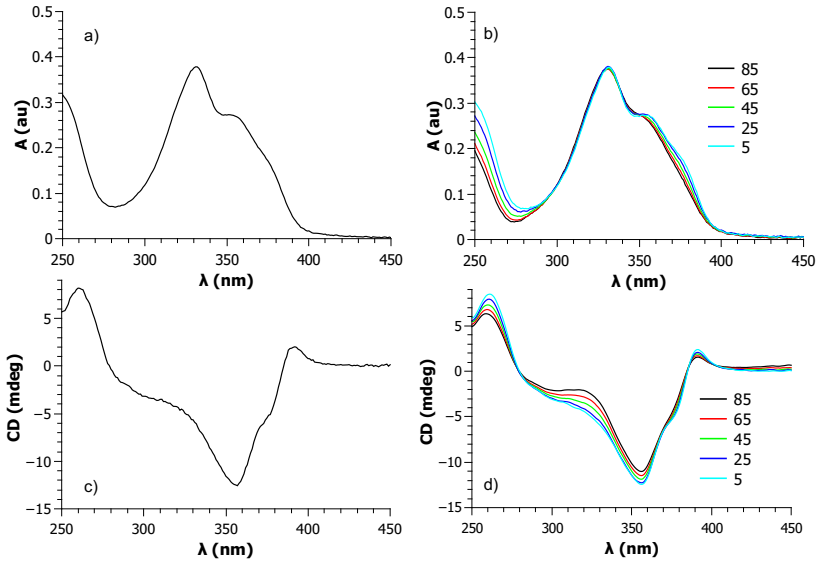


Figure 5.14. a) UV-Vis spectrum of **Cat IV *RSRR*** at 23 °C (296 K), b) VT UV-Vis spectra of **Cat IV *RSRR*** between 5 – 85 °C at the specified temperatures, c) CD spectrum of **Cat IV *RSRR*** at 23 °C and d) VT CD spectra of **Cat IV *RSRR***, between 5 – 85 °C at the specified temperatures. The solvent was a mixture of water-acetonitrile with 0.1% TFA.

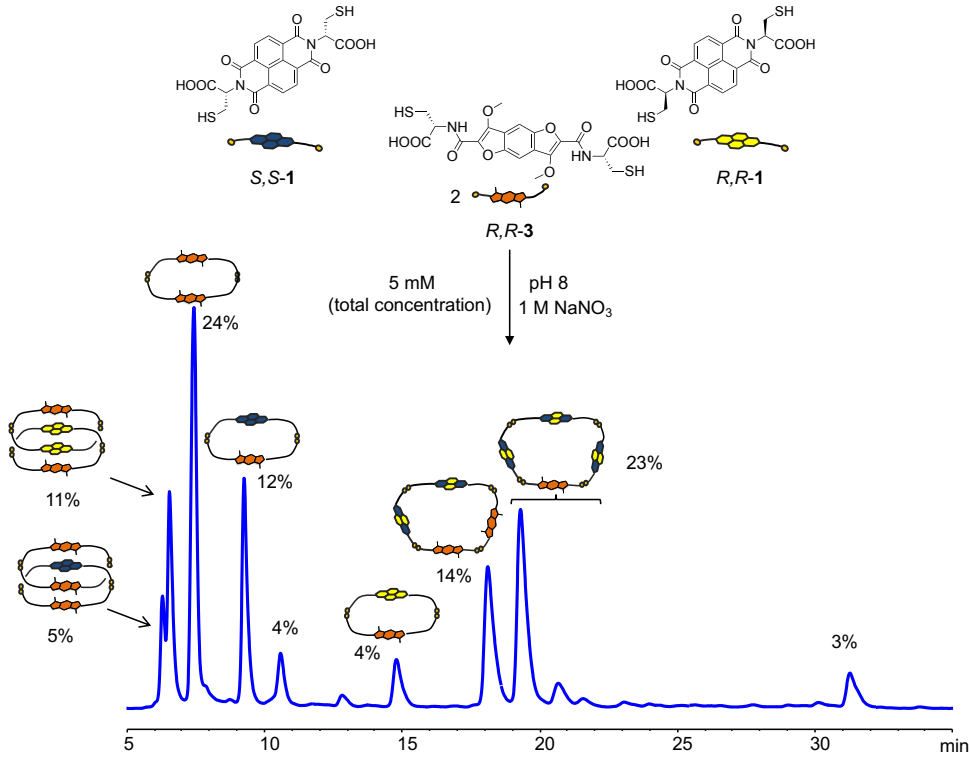


Figure 5.15. Reverse-phase HPLC analysis of *R,R*-**3**:*S,S*-**1**:*S,S*-**1** (2:1:1 molar ratio, 5 mM total concentration) library in the presence of 1 M $NaNO_3$. Absorbances recorded at 375 nm. The unlabelled peaks did not ionise and could not be identified.

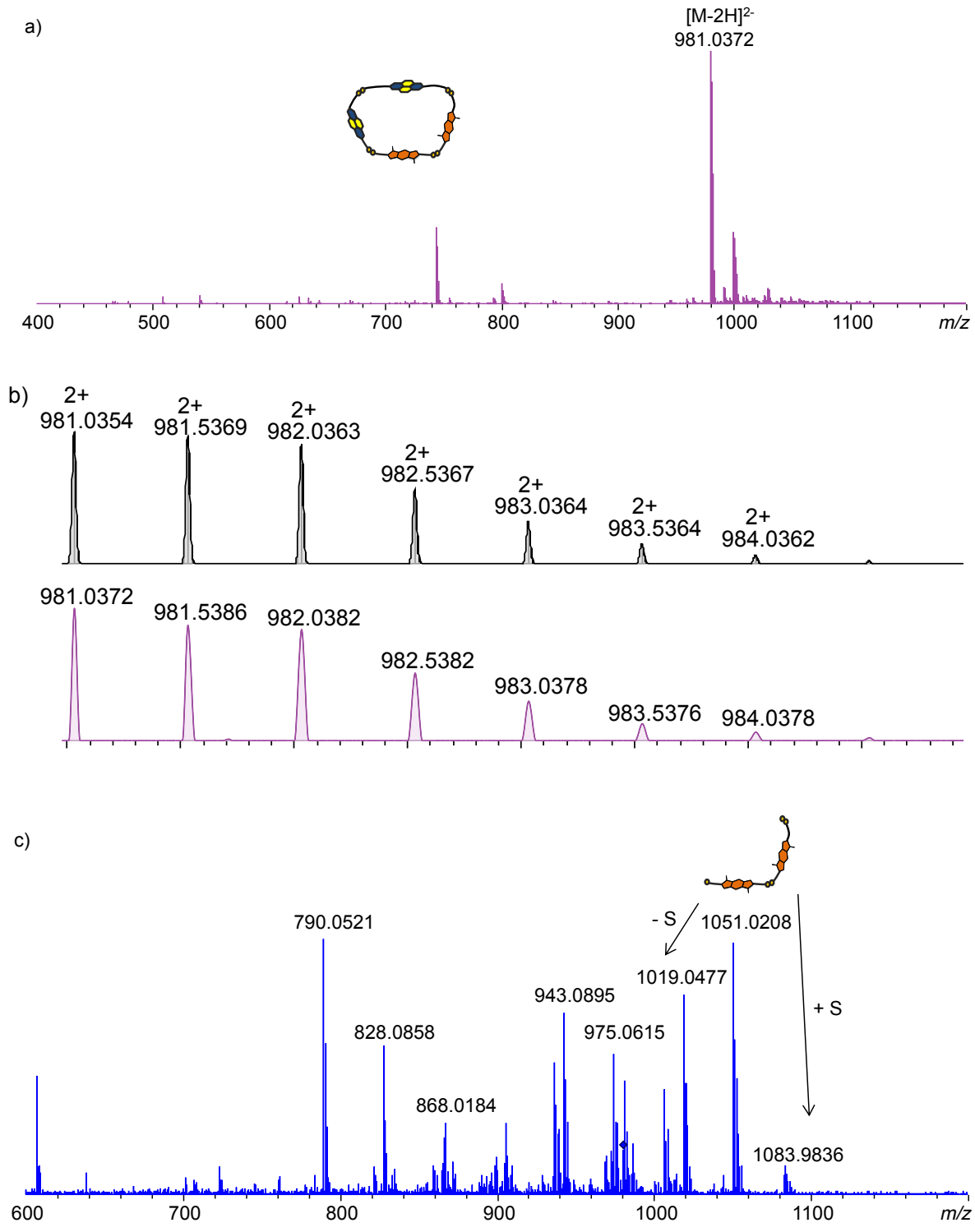


Figure 5.16. a) MS (*-ve*) of *R,R*-**3**, *S,S*-**1**, *R,R*-**1** tetramer, b) expansion of region of *R,R*-**3**, *S,S*-**1**, *R,R*-**1** tetramer and c) MS/MS (*-ve*) of *R,R*-**3**, *S,S*-**1**, *R,R*-**1** tetramer. The tetramer is based on two *R,R*-**3** units and two NDIs. It is impossible to know which enantiomer is employed (*S,S*-**1** or *R,R*-**1**), so the NDI cartoon representation contains both colours.

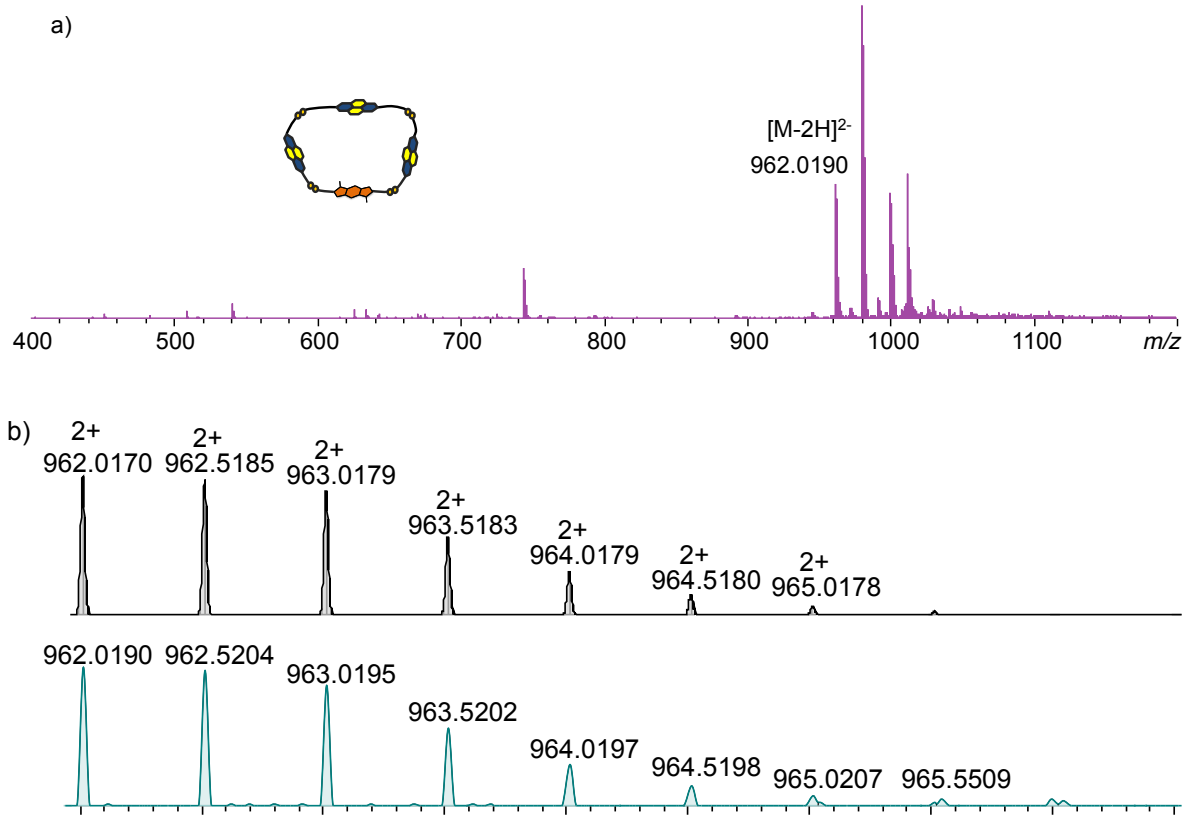
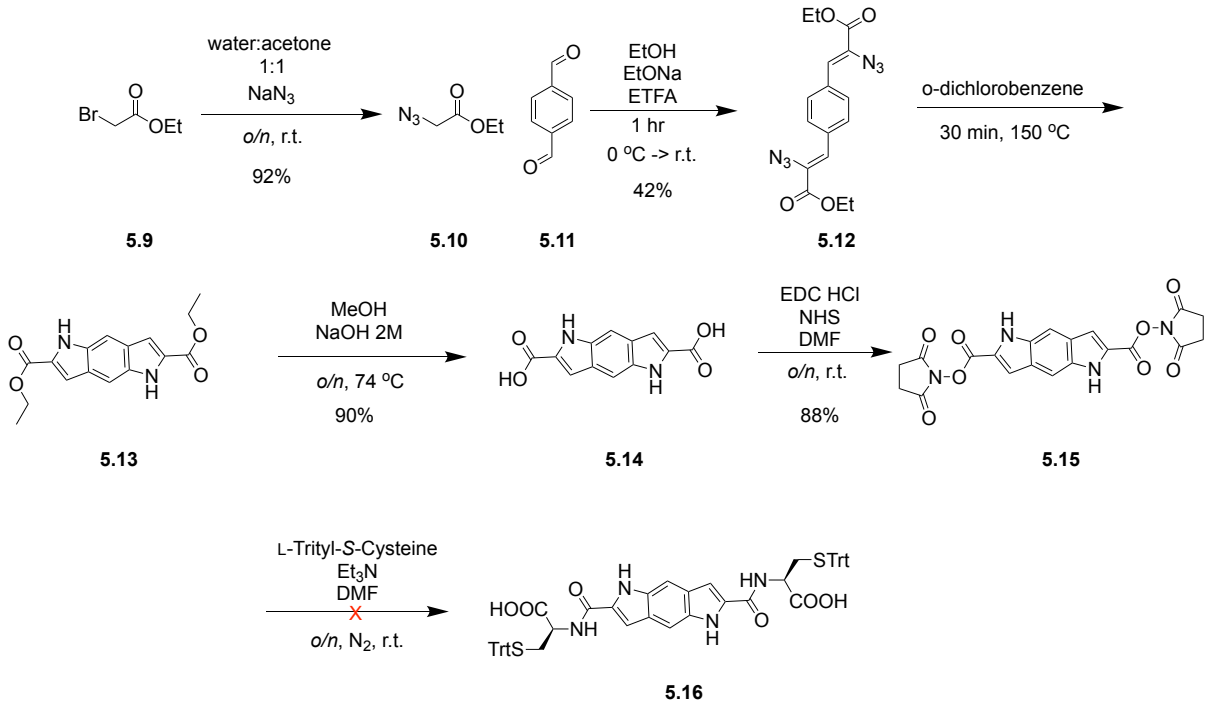


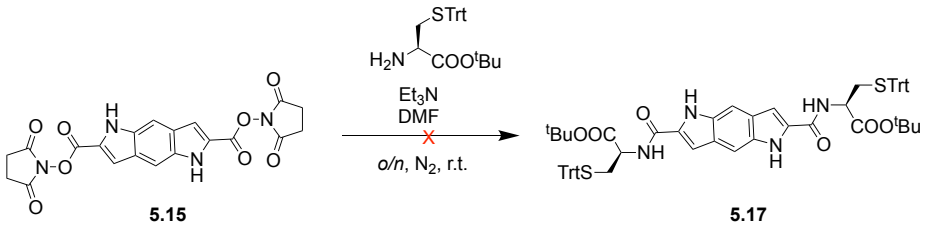
Figure 5.17. a) MS (*-ve*) of *R,R*-**3**, *S,S*-**1**, *R,R*-**1** tetramer and b) expansion of region of *R,R*-**3**, *S,S*-**1**, *R,R*-**1** tetramer. The tetramer is based on one *R,R*-**3** unit and three NDIs. It is impossible to know which enantiomer is employed (*S,S*-**1** or *R,R*-**1**), so the NDI cartoon representation contains both colours.

The general goal of my Ph.D. has been the design of new small molecules for disulphide-based DCC, building on the previous knowledge. Therefore, we have introduced electron-rich heteroaromatic building blocks. Besides sulphur and oxygen, we have also explored the influence of nitrogen as heteroatom in building block design. These molecules are generically called pyrroloindoles, a class of compounds that is similar to benzodithiophene and benzodifuran. Pyrroloindoles are inherent electron-rich molecules (the calculated Q_{zz} is -22.7 B), but should be less aromatic than benzodithiophenes.[2] The first attempt to make pyrroloindoles was based on a reported protocol; however, a key product in the synthetic route was obtained in low yield. The first step was the reaction between ethyl bromoacetate and sodium, giving ethyl azidoacetate. This was further used in a Knoevenagel type condensation with terephthaldehyde to yield compound **5.12**. The product obtained was subjected to the

Hemetsberger reaction, giving **5.13**. When looking at this molecule, one would expect to be soluble in chlorinated solvents. However, **5.13** was not soluble in $CHCl_3$, but only in DMSO. The ^1H NMR spectrum of this compound showed broad peaks in the aromatic region, which usually suggest polymerisation. Despite these issues, the product was used as it was in the next step (the yield is based on the product as well as the polymer – this is the reason why the yield is so high). The hydrolysis of the ester group, followed by the NHS activation of the carboxylic group, afforded the activated ester as in the case of the other compounds already described in the thesis (*i.e. R,R*-**2** and *R,R*-**3**). However, the next step, the reaction with *S*-trityl-L-cysteine, failed. ^1H NMR analysis showed more peaks than expected, suggesting side reactions (Scheme 5.2). Another attempt was to use the *tert*-butyl ester of *S*-trityl-L-cysteine (obtained by reacting *S*-trityl-L-cysteine with *tert*-butylacetate in acidic conditions), assuming that the carboxylic acid was the culprit. The reaction failed again (Scheme 5.3), so the strategy needed to be reassessed. An alternative synthetic pathway had to be designed.

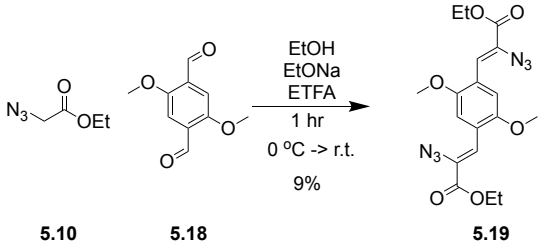


Scheme 5.2. The attempted synthetic scheme for compound **5.16**. The "X" means that the reaction did not work.

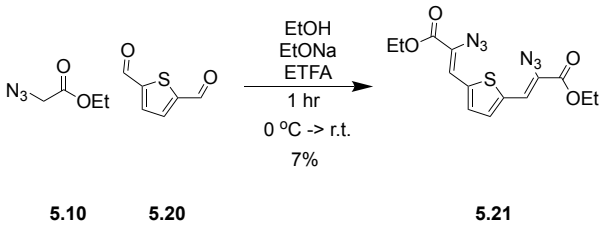


Scheme 5.3. Another attempt to put the cysteine on the PI core (compound **5.17**).

Other synthetic routes were designed towards the attachment of thiol functionality on the PI core. One way was to prepare the azido derivative with the cysteine attachment. Firstly, bromoacetic acid was reacted with sodium azide, following the NHS activation. The reaction with *S*-trityl-L-cysteine afforded the desired product. Unfortunately, the next step (*i.e.* Knoevenagel reaction) did not work. Therefore, the synthetic route was amended one more time. On the initial scheme (Scheme 5.2), other aldehydes were used instead of terephthaldehyde: 2,5-dimetoxyterephthaldehyde and thiophene-2,5-dicarbaldehyde (Schemes 5.4 and 5.5). This time, the Knoevenagel reaction worked in low yield. This suggested that such reaction works better with electron-poor molecules (because 2,5-dimethoxyterephthaldehyde is more electron-rich than terephthaldehyde).

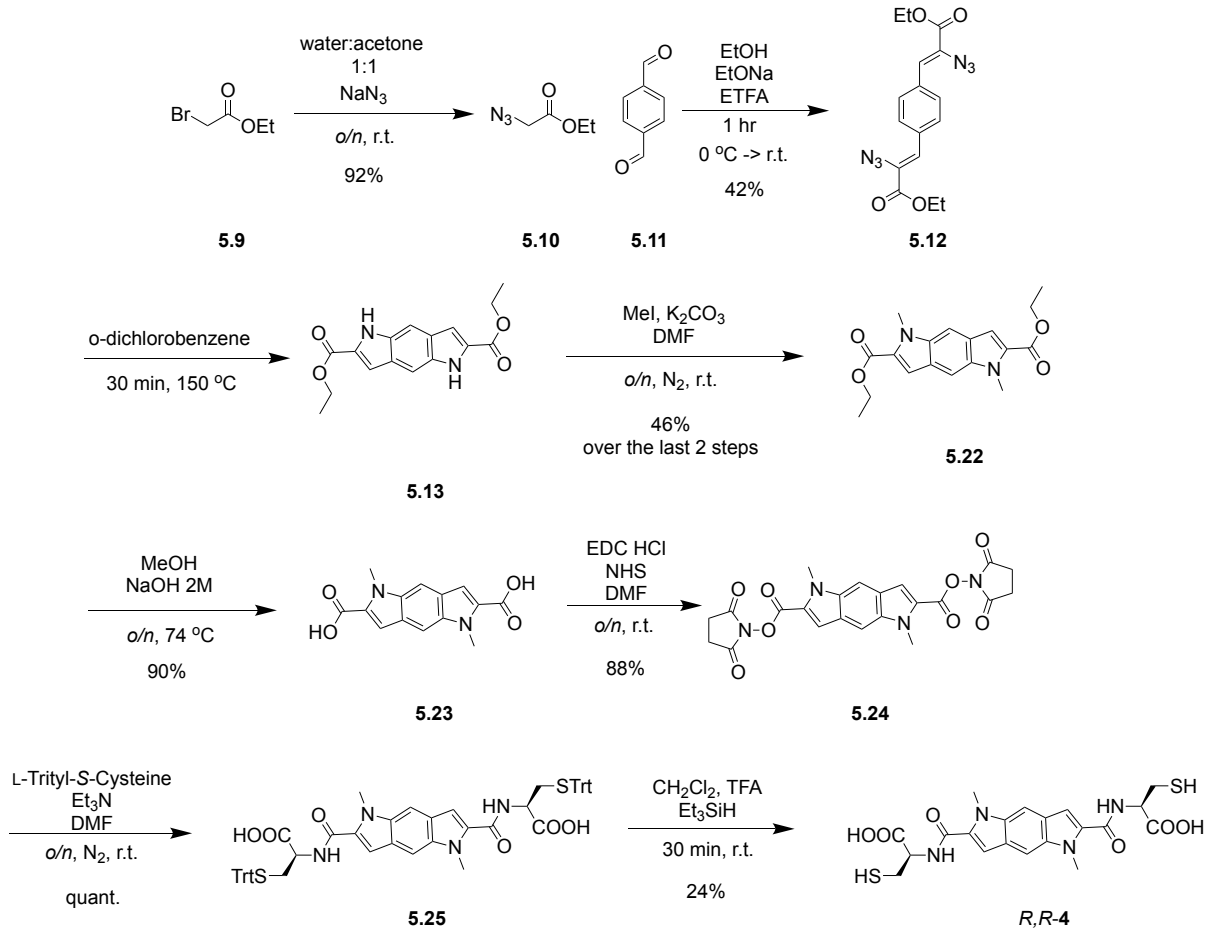


Scheme 5.4. The reaction for the synthesis of **5.19**.



Scheme 5.5. The reaction for the synthesis of **5.21**.

The synthetic route that finally worked involved the acylation of NH groups with methyl iodide, which afforded the compound **5.22**. NHS activation and reaction with cysteine allowed the synthesis of pyrroloindoles with thiol appendages (*R,R*-**4**). It seemed that the polymerisation problem was due to NH group. Another important aspect was that Hemetsberger reaction was performed at very dilutes scale to avoid the polymerisation as much as possible (Scheme 5.6).



Scheme 5.6. The synthetic scheme of *R,R*-**4**.

5.4. Analysis of *S,S*-1 and *R,R*-4 libraries

The building block *R,R*-**4** was dissolved in 10 mM NaOH and the pH was adjusted to 8. The DCLs were carried out in the presence and absence of 1 M $NaNO_3$. The distribution of the libraries in both cases showed the formation of the dimer as the major species (indicated by MS and MS/MS). *R,R*-**4** was also mixed with *R,R*-**1** with and without salt. The library without salt showed only the heterodimer, but when salt was added, the distribution was more diverse. Unfortunately, all the species were

macrocycles (dimer and tetramer); no interlocked structures were formed. Based on this observation, no heterochiral DCL was set up. Figures 5.18 – 5.22 show the MS characterisation of these species.

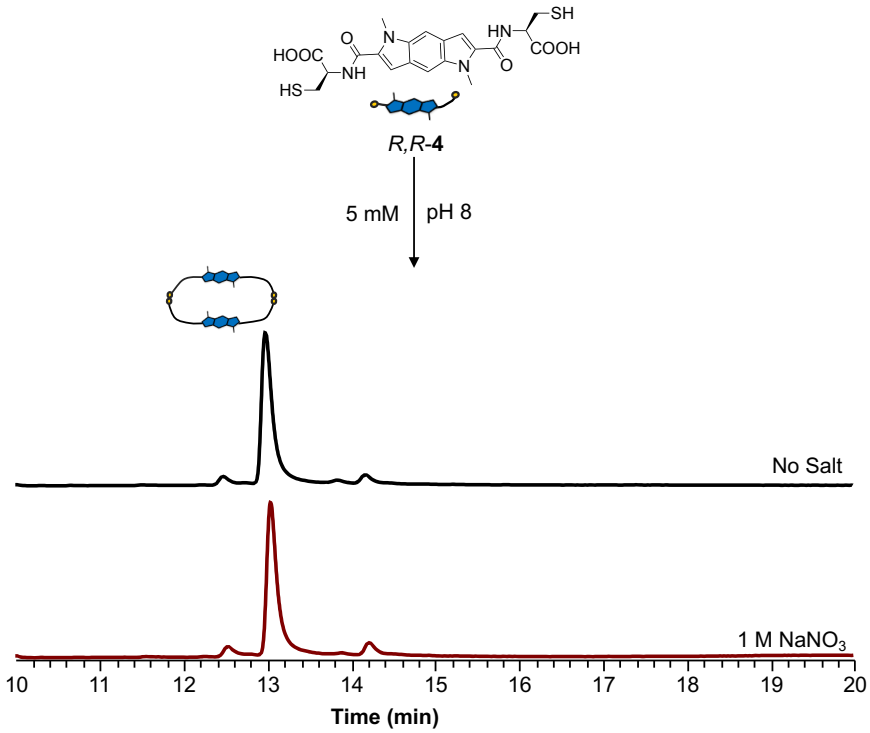


Figure 5.18. Reverse-phase HPLC analysis of *R,R*-**4** (5 mM total concentration) library without salt (top) and in the presence of 1 M $NaNO_3$ (bottom). Absorbances recorded at 325 nm. The unlabelled peaks did not ionise and could not be identified.

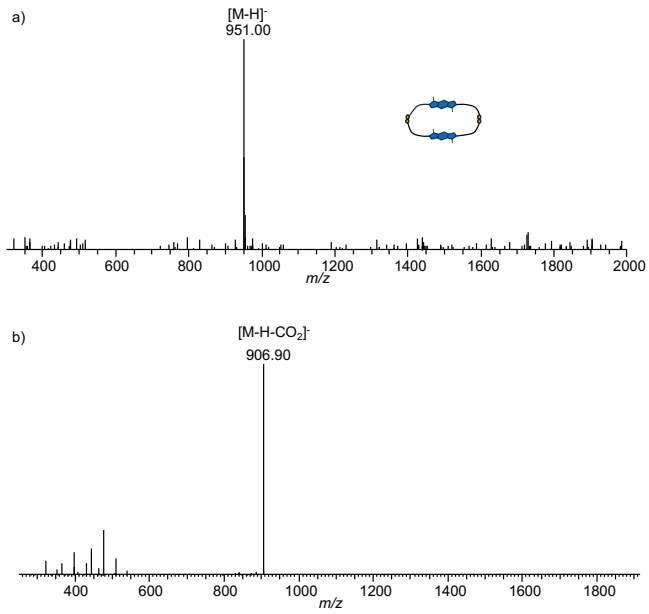


Figure 5.19. a) MS (*-ve*) of *R,R*-**4** dimer and b) MS/MS (*-ve*) of *R,R*-**4** dimer.

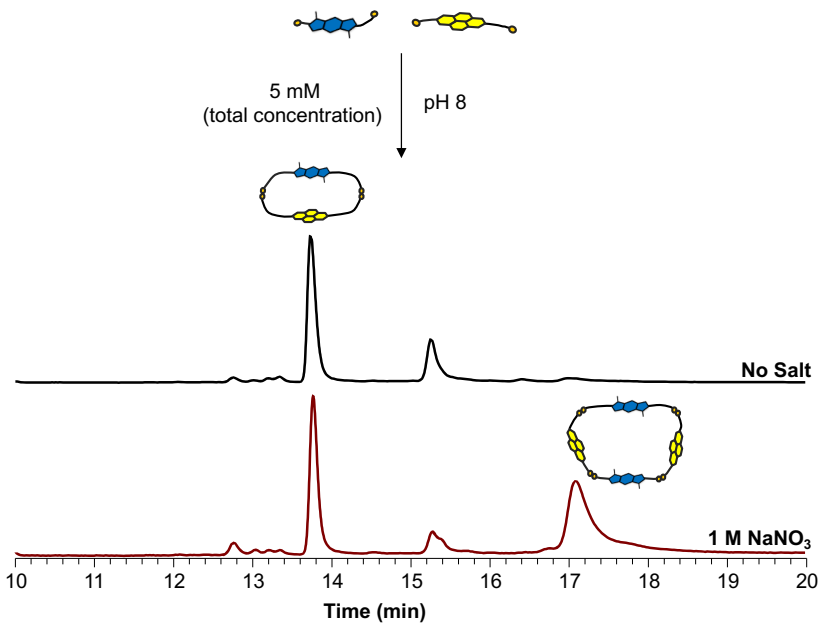


Figure 5.20. Reverse-phase HPLC analysis of *R,R*-**4**:*R,R*-**1** (1:1 molar ratio, 5 mM total concentration) library without salt (top) and in the presence of 1 M $NaNO_3$ (bottom). Absorbances recorded at 385 nm. The unlabelled peaks did not ionise and could not be identified.

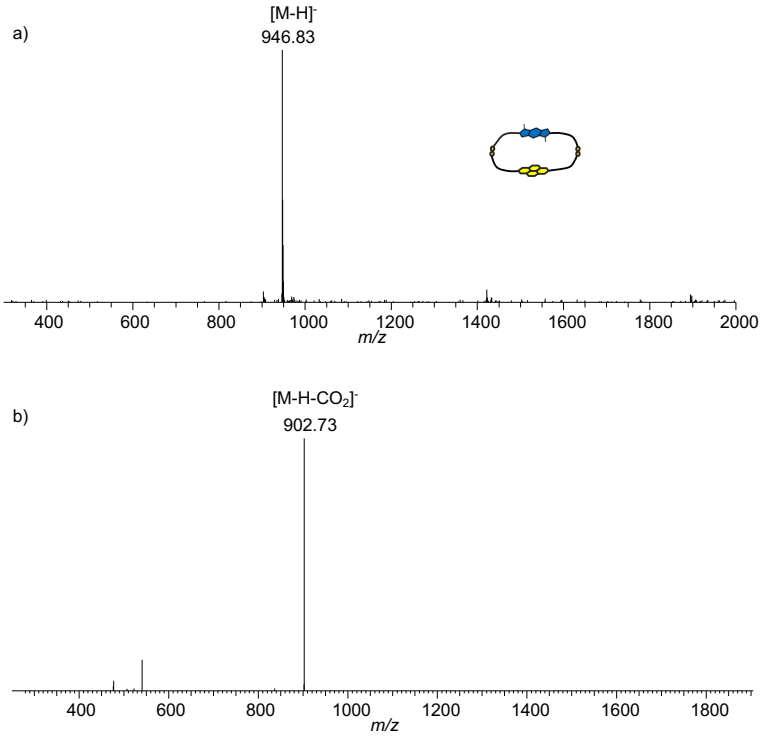


Figure 5.21. a) MS (*-ve*) of *R,R*-**4** and *R,R*-**1** heterodimer and b) MS/MS (*-ve*) of *R,R*-**4** and *R,R*-**1** heterodimer.

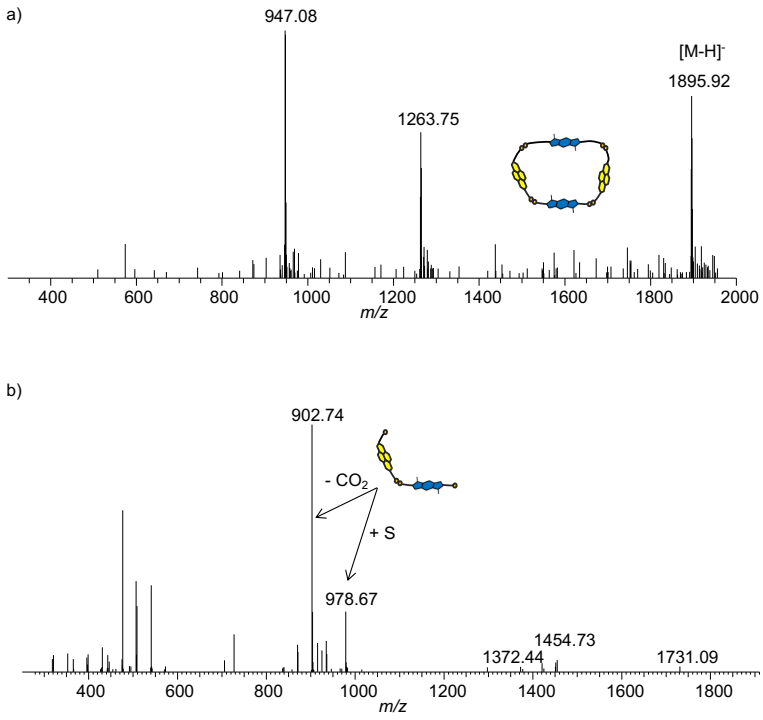


Figure 5.22. a) MS (*-ve*) of *R,R*-**4** and *R,R*-**1** tetramer and b) MS/MS (*-ve*) of *R,R*-**4** and *R,R*-**1** tetramer.

5.5 Conclusion

In conclusion, heteroatoms on the aromatic cores and peripheral substituents are of paramount importance for the outcome of DCLs based on this type of fused 5-6-5-membered rings aromatics. In the case of *R,R*-**3**, the methoxy groups along with the amide NH formed the pseudo-fusedpentacyclic structure,[3–5] which, in the presence of NDI *R,R*-**1** and *S,S*-**1** gives rise to two types of [2]catenanes. Therefore, *R,R*-**3** (BDF) behaves similarly to *R,R*-**2** (BDT), forming *DAAD* and *DADD* [2]catenanes in homochiral and heterochiral libraries, respectively. In the racemic DCL, the only catenanes formed in the *R,R*-**3** system are the same as in homo and heterochiral DCLs (*i.e. DAAD* and *DADD* [2]catenanes). The distribution of libraries of *R,R*-**2** and *R,R*-**3**, when set up individually, are different: *R,R*-**2** forms an all-donor [2]catenane, whereas the *R,R*-**3** does not. The structural divergence is more pronounced for *R,R*-**3** than for *R,R*-**2**. It is also more "clean" because the system is not as complex as in the case of *R,R*-**2**. The difference in behaviour between the DCL of *R,R*-**2** and that of *R,R*-**3** comes from the electron densities of furan and thiophene. Due to its high

electronegativity, oxygen's lone pair participates in conjugation with the π electrons in the ring, which makes BDT more aromatic than BDF.[2] Overall, BDT is more electron-rich than BDF based on Q_{zz} calculation. Therefore, one of the factors in the formation of the all-donor [2]catenane is the electronic nature of aromatic units forming it and the aromaticity itself. Another feature is the heteroatom present in the aromatic unit. In one case, sulphur is big and polarisable and more hydrophobic than oxygen. On the other case, the lone pair of oxygen would be available for hydrogen bonding.

In the case of pyrroloindole derivative, *R,R*-**4,** no interlocked structures have been observed. The lack of OMe groups and hence the lack of a pseudopentacyclic structure could be the reason for the markedly different library distribution. Further work on the synthetic pathway of *R,R*-**4** to introduce OMe, thus resemble *R,R*-**2** and *R,R*-**3** should be done. This work shows the power of DCC to distinguish between similar building blocks, and how the reactivity changes at small structural variations.

5.6. References

1 Z. Du, Y. Chen, W. Chen, S. Qiao, S. Wen, Q. Liu, D. Zhu, M. Sun and R. Yang, *Chem. - Asian J.*, 2014, **9**, 2621–2627.

2 K. E. Horner and P. B. Karadakov, *J. Org. Chem.*, 2013, **78**, 8037–8043.

3 P. Prabhakaran, A. Barnard, N. S. Murphy, C. A. Kilner, T. A. Edwards and A. J. Wilson, *Eur. J. Org. Chem.*, 2013, **2013**, 3504–3512.

4 J. C. Fuller, R. M. Jackson, T. A. Edwards, A. J. Wilson and M. R. Shirts, *PLoS ONE*, 2012, **7**, e43253.

5 J. P. Plante, T. Burnley, B. Malkova, M. E. Webb, S. L. Warriner, T. A. Edwards and A. J. Wilson, *Chem. Commun.*, 2009, 5091.

Chapter 6. Conclusions and Future Directions

The main finding of this thesis is the discovery of the first SDRRMs systems based on Dynamic Combinatorial Chemistry. Small variations in the building blocks used in DCLs give completely different library distributions and products. A small change of the heteroatom in the core of the building block (BDT *vs.* BDF with the NDI racemic mixture) leads to a cleaner SDRRM in the case of BDF, with the formation of two distinct catenanes. However, when the BDT is used as a racemic mixture, the divergence is not as pronounced and multiple species are formed. PI (building block with nitrogen) does not form the fused pseudo-pentacyclic structure as BDT and BDF, which results in a change of reactivity and DCL outcome (no interlocked molecules containing PI are observed). The use of another π-rich building block (DN) offers further insight into SDRRM. When the 1,5-DN is used, no SDRRM takes place, only a self-sorting process. However, in the case of the 2,6-DN, no species with mixed chirality is significantly formed.

The first ever aromatic all-donor catenane is reported to form in the presence of BDT and salt. In contrast, when iso-BDT, its isomer, is used no catenane is formed. This highlights that the geometry of the building block also plays an important role in the outcome of DCLs.

The design of a homochiral acceptor-donor-acceptor trimer and its use in DCL resulted in a 4_1 knot (from a 5 mM library), identified using a range of techniques: 1D and 2D NMR, CD, MS and computational data. Lowering the concentration from 5 to 0.5 mM, a [2]catenane is formed instead of the knot. When heterochiral trimer is used, the chirality mismatch influences the library distribution, with a [2]catenane formed at 5 mM concentration.

As future work, I suggest the synthesis of a pseudofused-pentacyclic PI for having the full picture of the influence of the most used heteroatoms in heterocycles in DCC. Another important direction is to synthesise trimers with the other building blocks (apart from BDT) as the first trimer gave a successful outcome. The use of flexible DN is also desirable because it will inform about the influence of rigid / flexible units in the outcome of SDRRM systems. Another promising idea is to desymmetrise the donor-based building blocks reported herein, and so, to synthesise all-donor-based trimers.

Chapter 7. Experimental Section

Materials and characterisation. All reagents were purchased from commercial suppliers: Acros Organics, Alfa Aesar, Sigma Aldrich, TCI Europe, Fluorochem and used without further purification. ^{1}H and ^{13}C NMR spectra were recorded on 500 MHz Agilent Propulse or 500 MHz Bruker Avance II+ (^{1}H NMR spectra at 500 MHz, ^{13}C NMR spectra at 125 MHz) instruments. Chemical shifts (δ) are reported in parts per million (ppm). Coupling constants are reported in Hertz (Hz), and signal multiplicity is denoted as singlet (s), doublet (d), doublet of doublets (dd), doublet of doublets of doublets (ddd), doublet of doublets of doublets of doublets (dddd), triplet (t), quartet (q), dt (doublet of triplets), td (triplet of doublets), h (heptet), multiplet (m) and broad (br). All spectra were acquired at 25 °C, unless otherwise stated, and were referenced to the residual solvent peaks. The common solvent impurities present in small amounts in ^{1}H and ^{13}C NMR spectra were water, acetone, CH_2Cl_2 or DMF. COSY, NOESY and HSQC spectra for isolated samples were recorded on a Bruker AV III 500 MHz fitted with Prodigy (nitrogen-cooled) cryoprobe. The microwave reactions were carried out in either CEM Discover or CEM Explorer 12 instruments. LC-MS studies were carried out on a Thermo Surveyor PDA Plus LC and LCQ classic ESI MS. Data was processed using the XCalibur software. Other LC studies (for *R,R*-**3** and for ***RSSR-NBN*** (0.5 mM) and ***SSSS*-NBN** (5 mM)) were done on HP1050 instrument equipped with a DAD lamp. The semi-preparative separations were performed on a HP1050 instrument. The individual HPLC / LC-MS methods at reported the end of each subchapter. The LC CD data was acquired on a Jasco CD 2095 connected to a SS420x A/D converter. All the CD, absorption and fluorescence data was acquired on an Applied Photophysics Chirascan spectrophotometer equipped with a Peltier temperature controller. The HRMS were either acquired at the Swansea University or at the University of Bath using a Bruker MicrOTOF Electrospray Time-Of-Flight mass spectrometer (ESI-TOF) coupled to an Agilent High Performance Liquid Chromatography (HPLC). The HRMS MS and MS/MS for the library components described in Chapters 4 and 5 were done on a Bruker MaXis HD ESI-QTOF mass spectrometer for high mass accuracy, coupled to a Thermo Scientific Dionex Ultra High Performance Liquid Chromatography (UPLC) unit.

DCL Set-up. A 5 mM (or different concentration) library was prepared by dissolving the building block in 10 mM aqueous NaOH, followed by titration with 100 mM NaOH / 100 mM HCl to adjust the pH to 8. The DCL solutions were stirred in closed-capped vials and analysed by LC-MS. Preparative libraries (≥ 3 mL scale) were made using the same method as for the analytical libraries.

LC-MS Method (Thermo Surveyor PDA Plus LC and LCQ classic ESI MS**).** ESI-MS spectra (negative ion) were acquired with a drying temperature of 250 °C, spray current 0.5 μA, sheath gas flow of 40 arb, spray voltage set to 4.5 kV, capillary voltage 13 V, tube lens -15.0 V. The mass range was set from *m/z* 150-2000, the number of microscans in scan time was 5 and the maximum injection time was 150.0 ms.

LC-MS Method (Bruker MaXis HD ESI-QTOF). **For Chapter 4 – for HRMS data:** ESI-MS spectra (negative ion) were acquired with drying temperature of 320 °C, collision energy -4 eV, dry gas 12 L/min. The mass range was set from *m/z* 350-3500. **For Chapter 5 – for HRMS data:** ESI-MS spectra (negative ion) were acquired with drying temperature of 320 °C, collision energy -4 eV, dry gas 12 L/min. The mass range was set from *m/z* 350-3500.

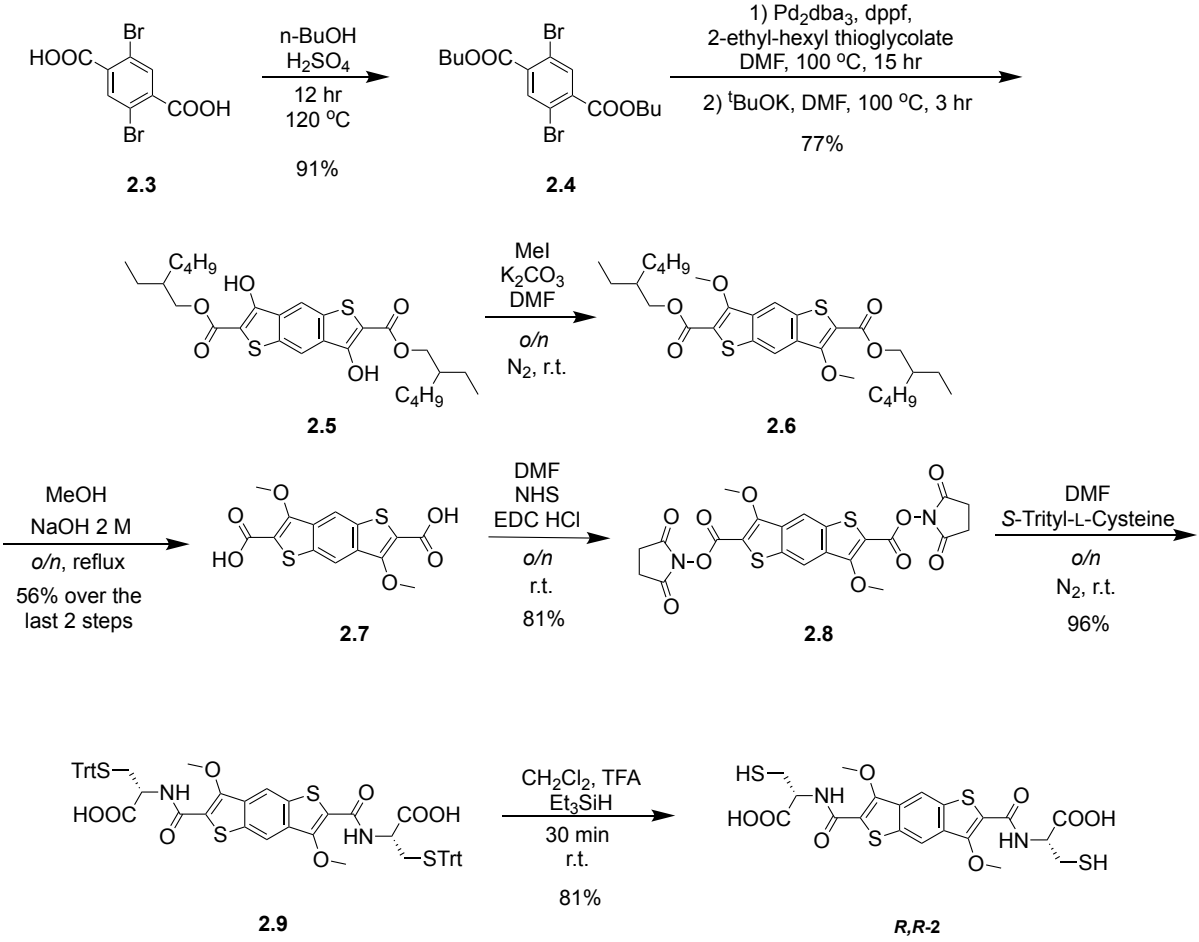
MS/MS Setting (Thermo Surveyor PDA Plus LC and LCQ classic ESI MS**).** Parent Mass (*m/z*): dependant on the species, Ionisation Width (*m/z*): 4.0, Normalized Collision Energy (%): 20, Activation Q: 0.250, Activation Time (msec) 30.0, the number of microscans in scan time was 5 and the maximum injection time was 150.0 ms. For Chapter 4 **MS/MS** for knot the resolution is 2 u and the intensity is lower.

MS/MS Setting (Bruker MaXis HD ESI-QTOF). **For Chapter 5 - for HRMS data:** Parent Mass (*m/z*): dependant on the species, Ionisation Width (*m/z*): 20.0, Collision Energy: 20 eV, drying temperature 240 °C, dry gas 12 L/min.

Fluorescence analysis: Fluorescence spectra were generally collected for samples with an absorbance below 0.1 au. PMU was set to 1000 V and bandwidth 4 nm. Time-per-point was 0.5 nm and excitation wavelength (for emission spectra) was set as the λ_{max} shown by the UV-Vis spectra. For the excitation spectra, the emission wavelength was set as the λ_{max} in Fluorescence emission spectra.

Computational studies. Geometry optimisation was performed using Avogadro[1] (Force field: UFF, Algorithm: Conjugate gradients). This was followed by MOPAC 2016[2] (Version 18.117M) PM7 semi-empirical optimisation using the COSMO water model with a 0.1 convergence factor, and Gabedit[3] 2.5.0 as interface. For CD prediction, ArgusLab version 4.0.1 (ZINDO-RPA, 80 excitations) was used.

Methods and synthesis for Chapter 2



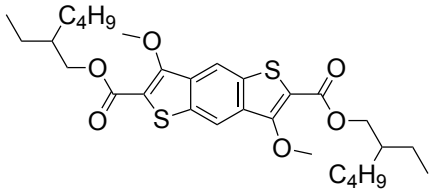
Compound **2.4** was synthesized following the published procedures.[4]

Compound **2.5** was synthesised using a modified procedure:[4]

2.4 (5.51 g, 12.70 mmoles, 1 equiv), $Pd_2(dba)_3$ (0.578 mg, 0.63 mmoles, 0.05 equiv) and dppf (0.70 mg, 1.26 mmoles, 0.1 equiv) were dissolved in dry DMF (70 mL) under N_2 atmosphere. To this mixture, $^{i}Pr_2NEt$ (11 mL, 63.14 mmoles, 4.97 equiv) and 2-ethylhexyl thioglycolate (6.34 mL, 30.16 mmoles, 2.37 equiv) were added and the

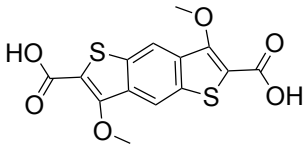
resulting solution was heated at 100 °C for 15 h. After cooling to r.t., ^{t}BuOK (4.10 g, 36.53 mmoles, 2.87 equiv) and DMF (40 mL) were added and the solution reheated at 100 °C for further 3 h. After cooling to r.t., 1 M HCl was added to acidify the solution, followed by extraction with toluene, filtration over Celite® and washings with ethyl acetate. After removing the solvent under reduced pressure, the residue was purified by column chromatography (CH_2Cl_2 : Petroleum Ether 60:40 *v*/*v*), giving compound **2.5** (5.19 g, 9.71 mmoles, 77%). The ^{1}H NMR spectrum is matching the one reported in the literature.[4]

Synthesis of **2.6**:



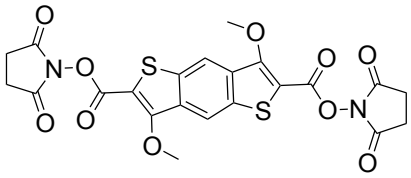
In a round-bottom flask, **2.5** (3.00 g, 5.60 mmoles,1 equiv) was dissolved in dry DMF (200 mL) under N_2 atmosphere, followed by addition of K_2CO_3 (3.88 g, 28.68 mmoles, 5 equiv). After stirring half an hour at r.t., MeI (1.4 mL, 22.48 mmoles, 4 equiv) was added and the reaction mixture was further stirred overnight. Water was added, and the compound was extracted with CH_2Cl_2. The organic phase was dried over $MgSO_4$ anhydrous, and the solvent removed *in vacuo* to yield an oily liquid, which was dried under high pressure. At this point the compound was impure, but it was used like this; no yield was determined at this stage. ^{1}H NMR (500 MHz, $CDCl_3$): δ 8.26 (s, 2H), 4.27 (dd, *J* = 5.7, 3.6 Hz, 4H), 4.20 (s, 6H), 1.73 (h, *J* = 5.9 Hz, 4H). The rest of the protons cannot be assigned because of the impurities. ^{13}C NMR (125 MHz): δ 169.6, 155.5, 140.0, 134.5, 134.3, 122.7, 117.1, 68.0, 67.8, 62.8, 38.9, 38.9, 36.4, 31.4, 30.5(1), 30.4(7), 29.0, 28.9, 23.9(3), 23.9(0). The ^{13}C NMR spectrum showed more peaks than expected because of the impurities. TOF MS ASAP+: *m/z* calcd for $C_{30}H_{42}O_6S_2$ $[M+H]^+$ 563.2501, found 563.2504.

Synthesis of **2.7**:



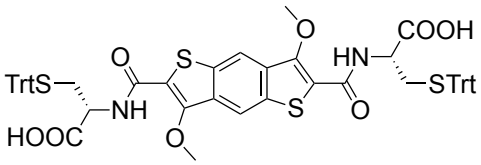
In a round-bottom flask, **2.6** (3.95 g) was dissolved in acetone (5 mL), followed by addition of MeOH (500 mL) and 2 M NaOH (250 mL). The reaction mixture was stirred overnight at reflux. Upon cooling, concentrated HCl was added until the pH became acidic and a yellow solid precipitated. The precipitate was filtered and washed with CH_2Cl_2 and acetone. The precipitate was collected and dried to yield **2.7** as a yellow solid (1.05 g, 3.12 mmoles, 58% over 2 steps). ^{1}H NMR (500 MHz, DMSO-d_6): δ 13.46 (br, 2H), 8.50 (s, 2H), 4.09 (s, 6H). ^{13}C NMR (125 MHz): δ 162.9, 154.8, 134.5, 134.0, 118.8, 117.9, 63.0. TOF MS ASAP+: *m/z* calcd for $C_{14}H_{10}O_6S_2$ $[M+H]^+$ 338.9997, found 338.9996.

Synthesis of **2.8**:



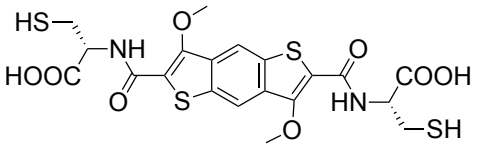
Compound **2.7** (1.05 g, 3.10 mmoles, 1 equiv) and *N*-hydroxysuccinimide (1.44 g, 12.50 mmoles, 4 equiv) were dissolved in dry DMF (110 mL) and cooled to 0 °C. EDC·HCl (2.41 g, 12.50 mmoles, 4 equiv) was added, and the reaction mixture was stirred for 15 minutes in the melting ice bath. The ice bath was removed and the reaction mixture was further stirred overnight at r.t. The solvent was removed under reduced pressure and a small amount of acetone was added. The suspension formed was added dropwise to a vigorously stirred aqueous solution of 1 M HCl. The precipitate obtained was collected by vacuum filtration and dried under reduced pressure to obtain a brownish solid (1.40 g, 2.60 mmoles, 85%). ^{1}H NMR (500 MHz, DMSO-d_6): δ 8.86 (s, 2H), 4.21 (s, 6H), 2.89 (s, 8H). ^{13}C NMR (125 MHz): δ 170.6, 159.4, 157.3, 135.4, 134.1, 119.7, 110.6, 63.7, 26.0. FTMS +pNSI: *m/z* calcd for $C_{22}H_{16}N_2O_{10}S_2$ $[M+NH_4]^+$ 550.0585, found 550.0580.

Synthesis of **2.9**:



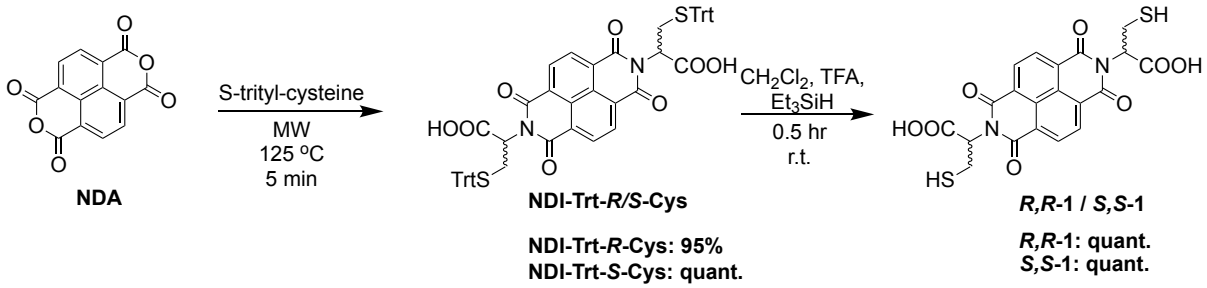
In a flame-dried round-bottom flask, *S*-trityl-L-cysteine (612.2 mg, 1.70 mmoles, 2.2 equiv) was added to a solution of **2.8** (401.8 mg, 0.76 mmoles, 1 equiv) in dry DMF (40 mL) under N_2 atmosphere. Et_3N (1 mL) was added, and the reaction mixture was stirred overnight at r.t. The solvent was removed under reduced pressure and a small volume of acetone was added to the mixture. This was dropwise precipitated into a vigorously stirred aqueous solution of 1 M HCl. The yellow precipitate was collected by filtration and dried (727.8 mg, 0.76 mmoles, quant.). ^{1}H NMR (500 MHz, DMSO-d_6): δ 13.19 (br, 2H), 8.70 (s, 2H), 8.42 (d, J = 7.6 Hz, 2H), 7.35-7.15 (m, 30H), 4.55 (td, J = 7.3, 4.6 Hz, 2H) 4.16 (s, 6H), 2.79 (dd, J = 12.3, 7.0 Hz, 2H). The other β protons of cysteine moiety are under the satellite of the solvent. ^{13}C NMR (125 MHz): δ 176.3, 165.4, 155.1, 149.3, 138.7, 137.2, 134.3, 133.2, 132.1, 71.3, 67.7, 56.5. FTMS -pNSI: *m/z* calcd for $C_{58}H_{48}N_2O_8S_4$ $[M-H]^-$ 1027.2221, found 1027.2198.

Synthesis of *R,R*-**2**:



In a round-bottom flask, **2.9** (500.4 mg, 0.49 mmoles) was dissolved in CH_2Cl_2 (10 mL) followed by addition of TFA (10 mL) and $SiEt_3H$ (1 mL). The reaction mixture was stirred 30 minutes at r.t. and the volatiles were subsequently removed. Et_2O was added, removed under pressure and the residue was re-dissolved in Et_2O to ensure the precipitation of the desired product. To this, *n*-hexane was added, the solid was filtered and further washed with *n*-hexane to yield *R,R*-**2** as an yellow solid that was dried (228.2 mg, 0.42 mmoles, 86%). ^{1}H NMR (500 MHz, DMSO-d_6): δ 13.29 (br, 2H), 8.69 (s, 2H), 8.47 (d, J = 7.4 Hz, 2H), 4.76 (dt, J = 7.4, 4.9, 2H) 4.20 (s, 6H). 3.15-3.00 (m, 4H). ^{13}C NMR (125 MHz): δ 171.6, 160.8, 150.5, 134.0, 132.6, 124.4, 118.2, 63.0,

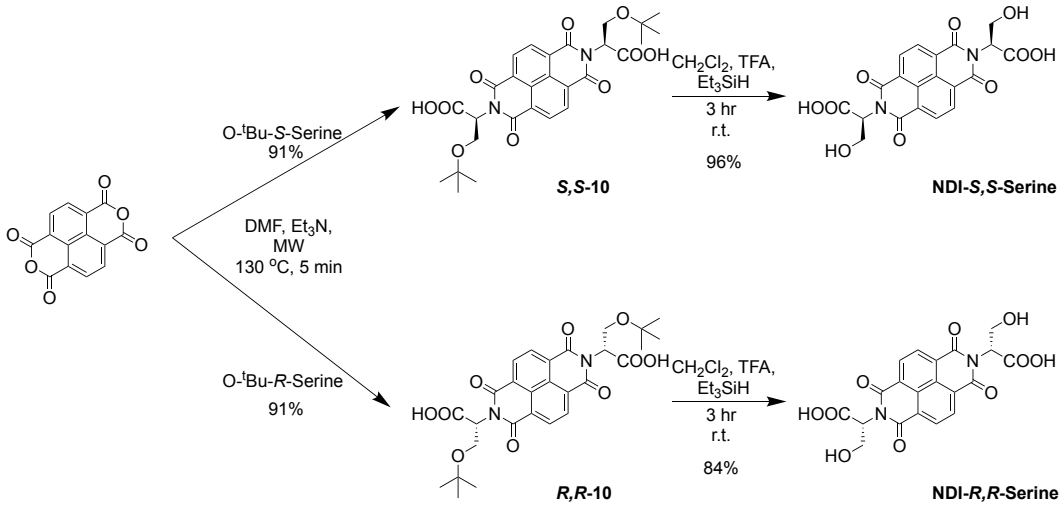
54.3, 26.2. FTMS -pNSI: *m/z* calcd for $C_{20}H_{19}N_2O_8S_4$ $[M-H]^-$ 543.0030, found 543.0032.



The building blocks *R,R*-**1** and **NDI-Trt-*R/S*-Cys** were synthesised following the previously published procedures.[5]

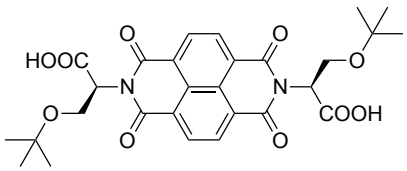
Synthesis of *S*,*S*-**1**:

NDI-Trt-*S*-Cys (201.0 mg, 0.21 mmoles, 1 equiv) was dissolved in a mixture of CH_2Cl_2 (3 mL) and TFA (3 mL), followed by the addition of Et_3SiH (0.3 mL). The reaction mixture was stirred for half an hour at r.t. and the volatiles were removed under reduced pressure. Et_2O was added, followed by the addition of *n*-hexane. The precipitate formed was filtered and dried (74.9 mg, 0.16 mmoles, 75%). ^{1}H NMR (500 MHz, DMSO-d_6) δ 13.22 (s, 2H), 8.78 (s, 4H), 5.73 (dd, J = 9.3, 5,3 Hz, 2H), 3.22 (dt, J = 14.2, 9.3 Hz, 2H), 2.72 (t, J = 8.9 Hz, 2H). The other two β protons are under the water peak. ^{13}C NMR (125 MHz) δ 170.0, 162.9, 131.8, 126.9, 126.5, 56.5, 23.2. TOF MS ASAP+: *m/z* calcd for $C_{20}H_{14}N_2O_8S_2$ $[M+H]^+$ 475.0270, found 475.0273.



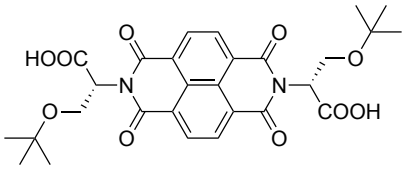
The building blocks *S,S*-**10** and *R,R*-**10** were synthesised using the published protocols (as **NDI-Trt-*S*-Cys**).[6]

Synthesis of *S,S*-**10**:



S,S-**10**: Started from NDA (202.6 mg, 0.76 mmoles, 1 equiv) and *O*-^{t}Bu-L-Serine (253.8 mg, 1.57 mmoles, 2.1 equiv). Obtained: (378.1 mg, 0.68 mmoles, 91%) *S,S*-**10**. ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.80 (s, 4H), 5.74 (t, *J* = 6.3 Hz, 2H), 4.43 (dd, *J* = 11.5, 6.3 Hz, 2H), 4.04 (dd, *J* = 11.6, 6.3 Hz, 2H), 1.46 (s, 18H). ^{13}C NMR (125 MHz) δ 167.6, 162.6, 131.4, 126.5, 83.2, 60.9, 55.7, 27.9. FTMS -pNSI: *m/z* calcd for $C_{28}H_{30}N_2O_{10}$: 553.1828 [*M*-H]$^{-}$; found 553.1829.

Synthesis of *R,R*-**10**:

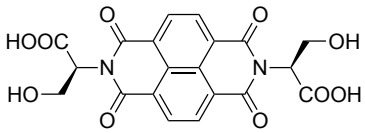


R,R-**10**: Started from NDA (200.7 mg, 0.75 mmoles, 1 equiv) and *O*-^{t}Bu-D-Serine (253.8 mg, 1.57 mmoles, 2.1 equiv). Obtained: (378.1 mg, 0.68 mmoles, 91%) *R,R*-**10**. ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.77 (s, 4H), 5.93 (t, *J* = 7.7 Hz, 2H), 4.35 (t, *J* = 8.6 Hz, 2H), 3.92 (dd, *J* = 9.1, 7.3 Hz, 2H), 1.24 (s, 18H). ^{13}C NMR (125 MHz) δ 162.3, 131.5, 126.4, 58.9, 52.9, 27.3. FTMS -pNSI: *m/z* calcd for $C_{28}H_{30}N_2O_{10}$: 553.1828 [*M*-H]$^{-}$; found 553.1828.

Deprotection procedure:

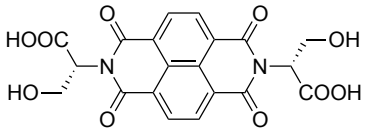
The protected species were dissolved in a mixture of CH_2Cl_2 (3 mL) and TFA (3 mL), followed by the addition of Et_3SiH (0.3 mL). The reaction mixture was stirred for 3 h at r.t. and the volatiles were removed under reduced pressure. Et_2O was added to the residue and the solid formed was filtered and dried.

Synthesis of **NDI-*S,S*-serine**:

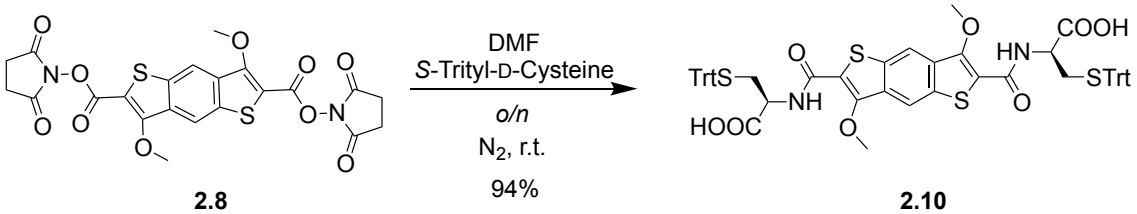


NDI-*S,S*-serine: Started from *S,S*-**10** (200.4 mg, 0.36 mmoles, 1 equiv) and formed **NDI-*S,S*-Serine** (152.6 mg, 0.35 mmoles, 96%). ^{1}H NMR (500 MHz, DMSO-d_6) δ 13.01 (br, 2H), 8.75 (s, 4H) 5.71 (t, J = 7.2 Hz, 2H), 4.90 (s, 2H), 4.09 (d, J = 7.2 Hz, 4H). ^{13}C NMR (125 MHz) δ 169.3, 162.5, 131.1, 126.1, 58.2, 55.6. FTMS -pNSI: *m/z* calcd for $C_{20}H_{14}N_2O_{10}$: 441.0576 [*M*-H]$^-$; found 445.0576.

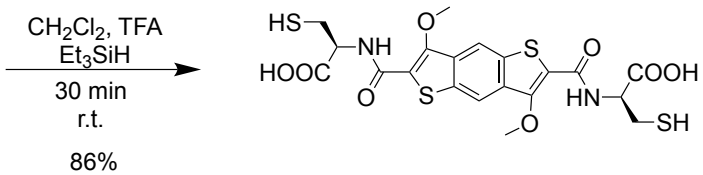
Synthesis of **NDI-*R,R*-serine**:



NDI-*R,R*-serine: Started from *R,R*-**10** (199.7 mg, 0.36 mmoles, 1 equiv) and formed **NDI-*R,R*-Serine** (133.3 mg, 0.30 mmoles, 84%). ^{1}H NMR (500 MHz, DMSO-d_6) δ 13.01 (br , 2H), 8.74 (s, 4H), 5.70 (t, J = 7.2 Hz, 2H), 4.83 (s, 2H*),* 3.98 (t, J = 7.2 Hz, 4H). ^{13}C NMR (125 MHz) δ 169.7, 162.9, 131.6, 126.6, 58.7, 56.0. FTMS -pNSI: *m/z* calcd for $C_{20}H_{14}N_2O_{10}$: 441.0576 [*M*-H]$^-$; found 445.0575.

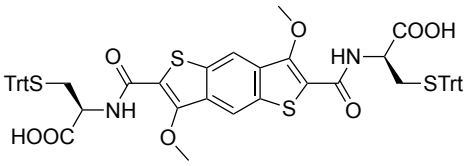


2.8 **2.10**



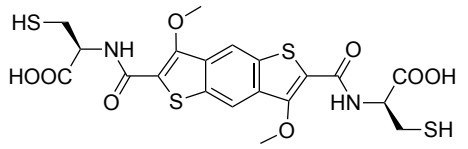
***S,S*-2**

Synthesis of **2.10**:



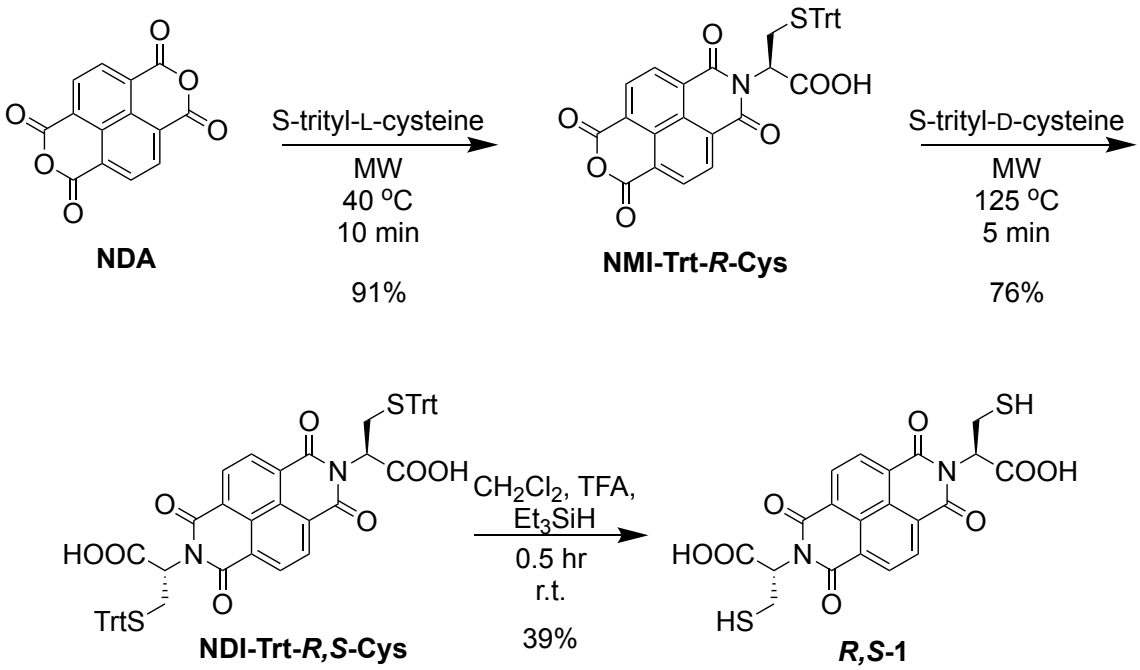
In a flame-dried round-bottom flask, *S*-trityl-D-cysteine (150.5 mg, 0.41 mmoles, 2.2 equiv) was added to a solution of **2.10** (100.6 mg, 0.19 mmoles, 1 equiv) in dry DMF (10 mL) under N_2 atmosphere. Dry Et_3N (0.5 mL) was added and the reaction was stirred overnight at r.t. The solvent was removed under reduced pressure and a small volume of acetone was added to the mixture. This was dropwise precipitated into a vigorously stirred aqueous solution 1 M HCl. The pale yellow precipitate was collected by filtration and dried (170.1 mg, 0.18 mmoles, 94%). ^{1}H NMR (500 MHz, DMSO-d_6): δ 13.20 (br, 2H), 8.71 (s, 2H), 8.43 (d, J = 7.2 Hz 2H), 7.40 – 7.15 (m, 30H), 4.56 (td, J = 7.2, 4.6 Hz, 2H), 4.17 (s, 6H), 2.80 (dd, J = 12.2, 7.1 Hz, 2H). The other β protons are under the satellite of DMSO peak. ^{13}C NMR (125 MHz, DMSO-d_6): δ 171.6, 160.7, 150.4, 144.6, 134.0, 132.6, 129.5, 128.5, 127.3, 124.0, 118.3, 66.6, 63.0, 51.7. FTMS +pNSI: *m/z* calcd for $C_{58}H_{48}N_2O_8S_4$ $[M+Na]^+$ 1051.2186, found 1051.2205.

Synthesis of *S,S*-**2**:



In a round-bottom flask, **2.10** (149.8 mg, 0.15 mmoles) was dissolved in CH_2Cl_2 (3 mL) followed by addition of TFA (3 mL) and $SiEt_3H$ (0.3 mL). The reaction mixture was stirred 30 minutes at r.t. and the volatiles were subsequently removed. Et_2O was added, removed under reduced pressure and the residue was re-dissolved in Et_2O to ensure the precipitation of the desired product. To this, *n*-hexane was added and the solid was filtered and further washed with *n*-hexane to yield *S,S*-**2** as an pale yellow precipitate that was dried (68.0 mg, 0.12 mmoles, 86%). ^{1}H NMR (500 MHz, DMSO-d_6): δ 13.30 (br, 2H), 8.70 (s, 2H), 8.48 (d, J = 7.4, 2H), 4.77 (dt, J = 7.4, 4.9 Hz, 2H), 4.20 (s, 6H), 3.15 – 3.01 (m, 4H). ^{13}C NMR (125 MHz, DMSO-d_6): δ 171.7,

160.8, 150.5, 134.0, 132.6, 124.4, 118.2, 63.0, 54.3, 26.3. FTMS +pNSI: *m/z* calcd for $C_{20}H_{19}N_2O_8S_4$ $[M+H]^+$ 545.0175, found 545.0168.

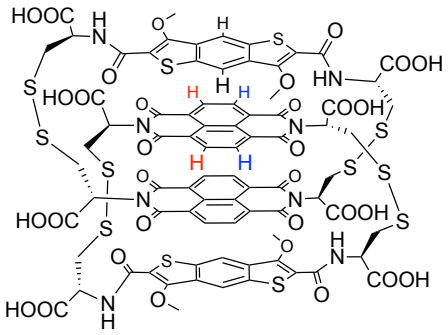


NMI-Trt-*R*-Cys was synthesised as described in the literature.[7]

NDI-Trt-*R,S*-Cys (318.6 mg, 0.33 mmoles) was dissolved in a mixture of CH_2Cl_2 (10 mL) and TFA (10 mL), followed by the addition of Et_3SiH (1 mL). The reaction mixture was stirred for 40 min at r.t. and the volatiles were removed under reduced pressure. Et_2O was added to the residue followed by *n*-hexane and the solid formed was filtered and dried to yield *R*,*S*-**1** (61.9 mg, 0.13 mmoles, 39%). ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.77 (s, 4H), 5.72 (dd, *J* = 9.3, 5.4 Hz, 2H), 3.36 (ddd, *J* = 14.0, 8.6, 5.3 Hz, 3H), 3.21 (dt, *J* = 14.1, 9.3 Hz, 3H), 2.71 (t, *J* = 9.0 Hz, 2H).^{13}C NMR (126 MHz, DMSO-d_6) δ 170.0, 162.8, 131.8, 126.9, 126.5, 56.5, 23.2. Q-TOF - ESI: *m/z* calcd for $C_{20}H_{24}N_2O_8S_2$ $[M-H]^+$ 473.0119, found 473.0114.

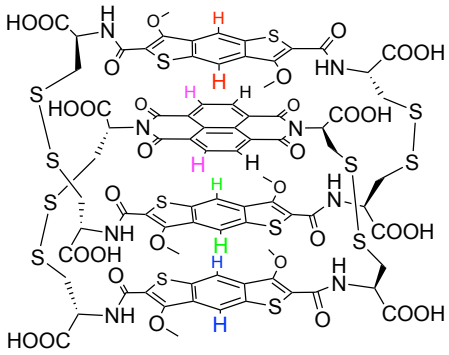
The synthesis of DNs was previously described.[8,9]

Cat I *RRRR* characterisation:



^{1}H NMR (500 MHz, D_2O): δ 9.24 (d, *J* = 5.6 Hz, 4H - NH), 8.33 (d, *J* = 7.5 Hz, 4H - H), 7.91 (d, *J* = 7.5 Hz, 4H - H), 7.06 (s, 4H - H), 5.66 (dd, *J* = 11.7, 3.4 Hz, 4H – α NDI protons), 4.17 (s, 12H), 3.96 – 3.86 (m, 8H – β protons), 3.74 – 3.64 (m, 9H – β protons). The other α protons next to BDT are under the solvent peak. ^{13}C NMR (from HSQC, D_2O): δ 130.9, 129.7, 115.3, 61.9, 56.1, 54.5, 43.9, 43.8, 40.1, 40.0.

Cat II *RSRR* characterisation:



^{1}H NMR (500 MHz, D_2O): δ 8.79 (d, *J* = 7.4 Hz, 2H – NH), 8.76 (d, *J* = 7.5 Hz, 2H – NH), 8.61 (d, *J* = 10 Hz, 2H – NH), 7.65 (d, *J* = 7.2 Hz, 2H – H), 7.63 (s, 2H – H), 7.59 (d, *J* = 7.2 Hz, 2H – H), 7.44 (s, 2H – H), 6.97 (s, 2H – H), 5.77 (dd, *J* = 11.8, 3.6 Hz, 2H – NDI α protons), 4.07 (s, 6H – OCH_3), 4.03 (s, 12H – OCH_3), 4.01 – 3.32 (m, 16H, NDI and BDT β protons). ^{13}C NMR (from HSQC, D_2O): δ 128.9, 128.1, 116.9, 116.8, 115.0, 62.2, 62.1, 61.3, 55.7, 54.3, 54.2, 45.1, 42.7, 39.3.

HPLC methods:

Analytical HPLC method for DCLs of Cat I *RRRR* (5 mM): Column**:** Develosil C_{30}, 5 x 0.3 cm, 3 µm. Injection volume: 2 µL; Flow rate: 1 mL/min; Temperature: 39 °C; Run time: 7 min; Elution profile:

Method A:

Time / min	Water (0.1% Formic Acid - FA)	CH_3CN (0.1% FA)
0	80%	20%
7	40%	60%
9	80%	20%

Preparative HPLC method for DCLs of Cat I *RRRR* (5 mM): Column**:** Develosil C_{30}, 5 x 2 cm, 5 µm; Injection volume: 900 µL; Flow rate: 6 mL/min; Temperature: 39 °C; Run time: 34 min; Elution profile:

Method B:

Time / min	Water (0.1% FA)	CH_3CN (0.1% FA)
0	90%	10%
2	90%	10%
3	75%	25%
23	58%	42%
30	40%	60%
32	35%	65%
34	80%	20%

Analytical HPLC method for DCLs of Cat II *RSRR* (5 mM): Column**:** Develosil C_{30}, 5 x 0.3 cm, 3 µm. Injection volume: 2 µL; Flow rate: 1 mL/min; Temperature: 39 °C; Run time: 7 min; Elution profile:

Method C:

Time / min	Water (0.1% FA)	CH_3CN (0.1% FA)
0	75%	25%
7	40%	60%
9	75%	25%

Preparative HPLC method for DCLs of Cat II *RSRR* (5 mM): Column**:** Develosil C_{30}, 5 x 2 cm, 5 µm; Injection volume: 600 µL; Flow rate: 6 mL/min; Temperature: 39 °C; Run time: 24 min; Elution profile:

Method D:

Time / min	Water (0.1% FA)	CH_3CN (0.1% FA)
0	75%	25%
20	55%	45%
24	75%	25%

Analytical HPLC method for DCLs of all libraries with *R,R*-1/*S,S*-1 with *R,R*-2 and/or *S,S*-2 (5 mM): Column**:** Ultra BiPhenyl, 5 x 0.3 cm, 3 µm. Injection volume: 3 µL; Flow rate: 0.5 mL/min; Temperature: 39 °C; Run time: 43 min; Elution profile:

Method E:

Time / min	Water (0.1% FA)	CH_3CN (0.1% FA)
0	80%	20%
5	75%	25%
20	70%	30%
30	65%	35%
35	60%	40%
40	50%	50%
43	10%	90%

Analytical HPLC method for DCLs of all libraries with *R,S*-1 with *R,R*-2 (5 mM): Column**:** Kromasil 100-5C8, 25 x 0.21 cm, 5 µm. Injection volume: 2 µL; Flow rate: 0.3 mL/min; Temperature: 39 °C; Run time: 45 min; Elution profile:

Method F:

Time / min	Water (0.1% FA)	CH_3CN (0.1% FA)
0	95%	5%
15	80%	20%
30	70%	30%
45	45%	55%

Analytical HPLC method for DCLs of all libraries with *R,R*-**1**/*S,S*-**1** with *R,R*-**2,6-DN**/*S,S*-**2,6-DN (5 mM):** Column**:** Symmetry C8, 15 x 0.39 cm, 5 µm. Injection volume: 2 µL; Flow rate: 1 mL/min; Temperature: 45 °C; Run time: 45 min; Elution profile:

Method G:

Time / min	Water (0.1% FA)	MeOH (0.1% FA)
0	60%	40%
20	40%	60%
45	40%	60%

Analytical HPLC method for DCLs of all libraries with *R,R*-**1**/*S,S*-**1** with *R,R*-**2,6-DN**/*S,S*-**2,6-DN (5 mM):** Column**:** Symmetry C8, 25 x 0.46 cm, 5 µm. Injection volume: 2 µL; Flow rate: 0.35 mL/min; Temperature: 32.5 °C; Run time: 45 min; Elution profile:

Method H:

Time / min	Water (0.1% FA)	MeOH (0.1% FA)
0-45 **(Isocratic)**	35%	65%

UV-Vis titration of NDI-*R,R-serine and* NDI-*S,S*-serine into **Y,** respectively**. Y** was prepared by dissolving the *R,R*-**2** building block in water at pH 8 and allowed to equilibrate over two days, yielding **Y** homodimer as the major product (more than 90%). From this stock solution, two solutions (each of 3.57 mM) were prepared: one containing the homodimer and 1 M $NaNO_3$ (soln. T) and the other containing either NDI-*R,R*-serine or NDI-*S,S*-serine (35.7 mM, 10 equiv) and 1 M $NaNO_3$ (soln. U). Small aliquots of soln. U were gradually added into soln. T, having 0.1, 0.3, 0.5, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2, 1.5, 2.0, 3.0, 4.0 and 5.0 equiv of NDI-serine at each titration point in the final solution.

Testing reversibility with DTT. For testing the reversible behaviour of the reactions, both Cat I *RRRR* and Cat II *RSRR* were subjected to dithiothreitol (DTT) after isolation. DTT is a dithiol-based compound that can restart the disulphide exchange and reform the library.

Circular Dichroism settings: Cat I *RRRR* and **Cat II *RSRR*** have a CD band

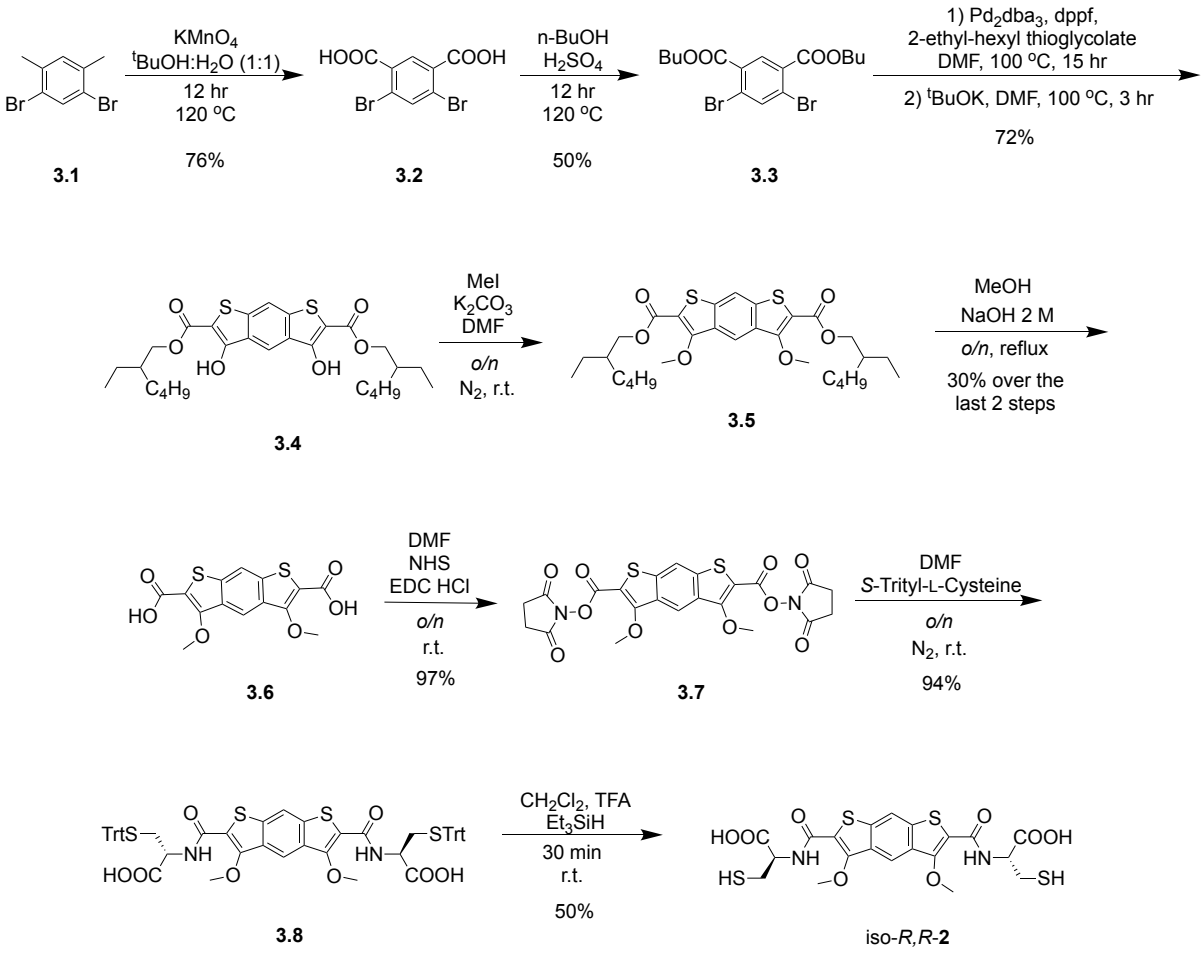
across the 550 nm region, which is expected for *D-A* interaction. The experiments were performed in a 1 or 10 mm pathlength cuvettes depending on the concentration, with a 4 nm bandwidth. Scan mode: 1 point/nm, time-per-point 0.5 s. The VT experiments (10 °C increment and setting time 300 s) were performed in the same cuvette using a stirrer bar.

Crystal data and structure refinement for 2.9

Empirical formula	$C_{58}H_{48}N_2O_8S_4 \cdot C_3H_5F_5O_2$
Formula weight	1189.29
Temperature	150.00(10) K
Wavelength	1.54184 Å
Crystal system	Orthorhombic
Space group	$P2_12_12_1$
Unit cell dimensions	a = 12.64620(10) Å α = 90°.
	b = 17.88170(10) Å β = 90°.
	c = 25.2987(2) Å γ = 90°.
Volume	5720.94(7) Å^3
Z	4
Density (calculated)	1.381 Mg/m^3
Absorption coefficient	2.172 mm^{-1}
F(000)	2472
Crystal size	0.180 x 0.180 x 0.100 mm^3
Theta range for data collection	3.026 to 73.054°.
Index ranges	-15<=h<=15, -19<=k<=22, -24<=l<=31
Reflections collected	46206
Independent reflections	11343 [R(int) = 0.0423]
Completeness to theta = 67.684°	99.8 %
Absorption correction	Semi-empirical from equivalents
Max. and min. transmission	1.00000 and 0.95696
Refinement method	Full-matrix least-squares on F^2
Data / restraints / parameters	11343 / 40 / 832
Goodness-of-fit on F^2	1.034

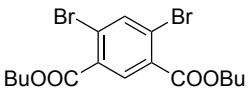
Final R indices [I>2sigma(I)]	R_1 = 0.0378, wR_2 = 0.0962
R indices (all data)	R_1 = 0.0404, wR_2 = 0.0984
Absolute structure parameter	0.007(5)
Largest diff. peak and hole	0.486 and -0.345 e·Å^{-3}

Methods and synthesis for Chapter 3



Compound **3.2** was synthesized following the published procedures.[10]

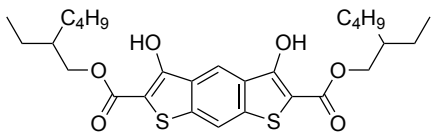
Synthesis of **3.3**:



In a round-bottom flask, **3.2** (5.57 g, 17.2 mmoles, 1 equiv), *n*-BuOH (81 mL) and conc. H_2SO_4 (1.2 mL) were mixed and refluxed overnight. After cooling the reaction mixture, most of *n*-BuOH was removed *in vacuo* followed by neutralisation

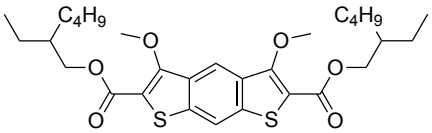
with sat. $NaHCO_3$. The solution was extracted with ethyl acetate, and the organic layer was washed with water then dried over $MgSO_4$. The solvent was removed under reduced pressure and the compound was purified by silica-gel column chromatography (eluents: petroleum ether, followed by CH_2Cl_2). The desired fractions were collected and the solvent was evaporated, yielding **3.3** as a white compound (3.75 g, 8.61 mmoles, 50%). ^{1}H NMR (500 MHz, $CDCl_3$): δ 8.25 (s, 1H), 8.02 (s, 1H), 4.36 (t, *J* = 6.6 Hz, 4H), 1.76 (m, 4H), 1.48 (m, 4H), 0.99 (t, *J* = 7.4 Hz, 6H). ^{13}C NMR (125 MHz, $CDCl_3$): 164.7, 139.7, 133.8, 131.3, 125.5, 66.0, 30.5, 19.2, 13.7. FTMS +pNSI: *m/z* calcd for $C_{16}H_{20}Br_2O_4$ $[M+H]^+$ 434.9801, found 434.9791.

Synthesis of **3.4**:



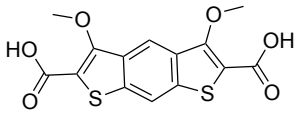
In a round-bottom flask, **3.3** (3.75 g, 8.61 mmoles, 1 equiv), $Pd_2(dba)_3$ (0.39 mg, 0.43 mmoles, 0.05 equiv) and dppf (0.48 mg, 0.86 mmoles, 0.1 equiv) were dissolved in DMF (48 mL) under N_2 atmosphere. To this mixture, ^{i}Pr$_2$NEt (7.4 mL, 42.19 mmoles, 4.9 equiv) and 2-ethylhexyl thioglycolate (4.60 mL, 21.52 mmoles, 2.50 equiv) were added, and the resulting solution was heated at 100 ^{o}C for 15 h. After cooling to r.t., ^{t}BuOK (2.80 g, 24.97 mmoles, 2.90 equiv) and DMF (20.2 mL) were added and the solution reheated to 100 ^{o}C for further 3 h. After cooling again to r.t., 1 M HCl was added to acidify the solution, followed by extraction with toluene, filtration over Celite® and washings with ethyl acetate. After removing the solvent, the compound was purified by column chromatography (CH_2Cl_2 : petroleum ether 20:80 *v*/*v*, then CH_2Cl_2 : petroleum ether 80:20 *v*/*v*), giving compound **3.4** (3.29 g, 6.16 mmoles, 72%). ^{1}H NMR (500 MHz, $CDCl_3$): δ 10.25 (s, 2H), 8.55 (s, 1H), 8.06 (s, 1H), 4.30 (dd, *J* = 5.8, 2.9 Hz, 4H), 1.74 (p, *J* = 6.2 Hz, 2H), 1.51 – 1.28 (m, 18H), 1.00 – 0.80 (m, 18H). ^{13}C NMR (125 MHz, $CDCl_3$) δ 167.15, 139.10, 128.56, 117.74, 116.97, 67.87, 38.84, 30.45, 28.93, 23.89, 22.94, 14.02, 11.07. +p CI Full MS: *m/z* calcd for $C_{28}H_{38}O_6S_2$ $[M+NH_4]^+$ 552.2, found 552.2.

Synthesis of **3.5**:



In a round-bottom flask, **3.4** (3.00 g, 5.60 mmoles,1 equiv) was dissolved in dry DMF (187 mL) under N_2 atmosphere followed by addition of K_2CO_3 (3.88 g, 28.68 mmoles, 5 equiv). After stirring half an hour at r.t., MeI (1.4 mL, 22.48 mmoles, 4 equiv) was added and further stirred overnight. Water was added, and the compound was extracted with CH_2Cl_2. The organic phase was dried over $MgSO_4$ anhydrous and the solvent removed *in vacuo* to yield an oily liquid, which was dried under high pressure. At this stage the compound was impure, but used as it was; no yield was determined at this stage. ^{1}H NMR (500 MHz, $CDCl_3$): δ 8.39 (d, *J* = 0.9 Hz, 1H), 8.08 (d, *J* = 0.9 Hz, 1H), 4.26 (dd, *J* = 5.7, 4.3 Hz, 4H), 4.23 (s, 6H), 1.72 (h, *J* = 6.1 Hz, 4H). The rest of the protons cannot be assigned because of the impurities' presence. ^{13}C NMR (125 MHz, $CDCl_3$): δ 169.49, 162.52, 161.72, 156.57, 145.69, 137.78, 132.27, 132.02, 122.60, 117.32, 116.87, 116.44, 114.91, 68.57, 67.83, 67.70, 63.07, 62.93, 61.96, 38.92, 38.60, 36.47, 31.44, 30.52, 30.24, 30.19, 28.97, 28.65, 23.94, 23.67, 22.99, 22.84, 20.72, 14.06, 13.88, 11.10, 10.89. There are more peaks in the ^{13}C NMR spectrum than expected because of the impurity. +p APCI corona: *m/z* calcd for $C_{30}H_{42}O_6S_2$ $[M+H]^+$ 563.2496, found 563.2489.

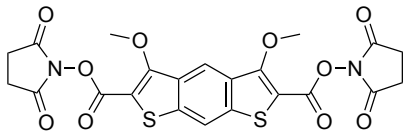
Synthesis of **3.6**:



In a round-bottom flask, **3.5** (6.65 g) was dissolved in acetone (5 mL), followed by addition of MeOH (400 mL) and 2 M NaOH (200 mL). The reaction mixture was stirred overnight at reflux. Upon cooling, concentrated HCl (32%) was added until the pH became acidic and a yellow solid precipitated. The precipitate was filtered off and washed with CH_2Cl_2 and acetone. The precipitate was collected and dried to yield **3.6** as a pale yellow solid (0.56 g, 1.66 mmoles, 30% over 2 steps). ^{1}H NMR (500 MHz,

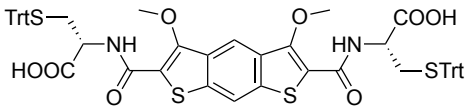
DMSO-d_6): δ 13.46 (br, 2H), 8.56 (s, 1H), 8.24 (s, 1H), 4.13 (s, 6H). ^{13}C NMR (125 MHz, DMSO-d_6): δ 162.9, 155.7, 137.0, 132.2, 118.0, 116.7, 116.5, 63.1. FTMS -pNSI: *m/z* calcd for $C_{14}H_{10}O_6S_2$ $[M-H]^+$ 336.9846, found 338.9843.

Synthesis of **3.7**:



In a round-bottom flask, compound **3.6** (0.56 g, 1.66 mmoles, 1 equiv) and *N*-hydroxysuccinimide (0.77 g, 6.67 mmoles, 4 equiv) were dissolved in dry DMF (56 mL) and cooled to 0 °C with an ice bath. EDC·HCl (1.03 g, 6.64 mmoles, 4 equiv) was added and the reaction was stirred for 15 minutes in the melting ice bath. The ice bath was removed, and the reaction was further stirred overnight at r.t. The solvent was removed under reduced pressure and a small amount of acetone was added. The suspension formed was dropwise precipitated into a vigorously stirred aqueous solution of 1 M HCl. The precipitate obtained was collected by vacuum filtration and dried under reduce pressure to obtain a pale yellow solid (0.86 g, 1.61 mmoles, 97%). ^{1}H NMR (500 MHz, DMSO-d_6): δ 8.82 (s, 1H), 8.57 (s, 1H), 4.25 (s, 6H), 2.89 (s, 8H). ^{13}C NMR (125 MHz, DMSO-d_6): δ 170.7, 160.6, 157.3, 139.2, 131.3, 119.3, 119.0, 107.8, 64.0, 26.0. FTMS +pAPCI corona: *m/z* calcd for $C_{22}H_{16}N_2O_{10}S_2$ $[M]^+$ 532.0241, found 532.0241.

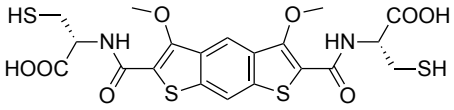
Synthesis of **3.8**:



In a flame-dried round-bottom flask, *S*-trityl-L-cysteine (527.8 mg, 1.45 mmoles, 2.2 equiv) was added to a solution of **3.8** (352.8 mg, 0.66 mmoles, 1 equiv) in dry DMF (30 mL) under N_2 atmosphere. Dry Et_3N (1.8 mL) was added and the reaction mixture was stirred overnight at r.t. The solvent was removed under reduced pressure and a

small volume of acetone was added to the mixture. This was dropwise precipitated into a vigorously stirred aqueous solution of 1 M HCl. The pale yellow precipitate was collected by filtration and dried (633.3 mg, 0.62 mmoles, 94%). ^{1}H NMR (500 MHz, DMSO-d_6): δ 13.18 (br, 2H), 8.70 (s, 1H), 8.48 – 8.37 (m, 3H), 7.34 – 7.15 (m, 30H), 4.55 (td, J = 7.2, 4.6 Hz, 2H), 4.20 (s, 6H), 2.79 (dd, J = 12.2, 7.0 Hz, 2H), 2.61 (dd, J = 12.2, 4.5 Hz, 2H). ^{13}C NMR (125 MHz, DMSO-d_6): δ 171.6, 160.7, 151.2, 144.6, 136.2, 130.9, 129.5, 129.5, 129.4, 128.5, 127.3, 122.4, 66.5, 63.1, 51.7. FTMS -pNSI: *m/z* calcd for $C_{58}H_{48}N_2O_8S_4$ $[M-2H]^{2-}$ 513.1074, found 513.1084.

Synthesis of iso-R,R-2:



In a round-bottom flask, **3.8** (124.1 mg, 0.12 mmoles) was dissolved in CH_2Cl_2 (2 mL) followed by addition of TFA (2 mL) and $SiEt_3H$ (0.2 mL). The reaction mixture was stirred 30 minutes at r.t., and the volatiles were subsequently removed under reduced pressure. Et_2O was added, removed under reduced pressure and the residue was re-dissolved in Et_2O to ensure the precipitation of the desired product. To this, *n*-hexane was added and the solid was filtered off and further washed with *n*-hexane to yield iso-R,R-2 as a pale yellow solid that was dried (32.5 mg, 0.06 mmoles, 50%). ^{1}H NMR (500 MHz, DMSO-d_6): δ 13.28 (br, 4H), 8.68 (s, 1H), 8.58 (s, 1H), 8.51 – 8.28 (m, 5H), 4.76 (dt, J = 7.4, 4.8 Hz, 4H), 4.27 – 4.08 (m, 12H), 3.16 – 2.98 (m, 8H). By the time the ^{1}H NMR spectrum was acquired, the oxidation of the thiols already started; hence the disulphide exchange started as well. Because of this, a ^{13}C NMR spectrum was not recorded. -p CI: *m/z* calcd for $C_{20}H_{19}N_2O_8S_4$ $[M-H]^{-}$ 543.0, found 543.0.

HPLC methods:

Analytical HPLC method for DCLs of *R,R*-2 (5 mM): Column**:** Kromasil 100-5C8, 25 x 0.21 cm, 5 µm. Injection volume: 2 µL; Flow rate: 0.3 mL/min; Temperature: 39 °C; Run time: 10 min; Elution profile:

Method I:

Time / min	Water (0.1% FA)	CH_3CN (0.1% FA)
0	70%	30%
10	50%	50%

Preparative HPLC method for DCLs of *DDDD* [2]catenane *RRRR*(5 mM): Column**:** Kromasil 100-5C8, 25 x 0.8 cm, 5 µm; Injection volume: 500 µL; Flow rate: 3.5 mL/min; Temperature: 39 °C; Run time: 12.5 min; Elution profile:

Method J:

Time / min	Water (0.1% FA)	CH_3CN (0.1% FA)
0	70%	30%
12.5	50%	50%

Analytical HPLC method for DCLs of *R,R*-2 with *S,S*-2 (5 mM): Column**:** Kromasil C8 5u, 25 x 0.46 cm, 5 µm. Injection volume: 4 µL; Flow rate: 1 mL/min; Temperature: 39 °C; Run time: 37 min; Elution profile:

Method K:

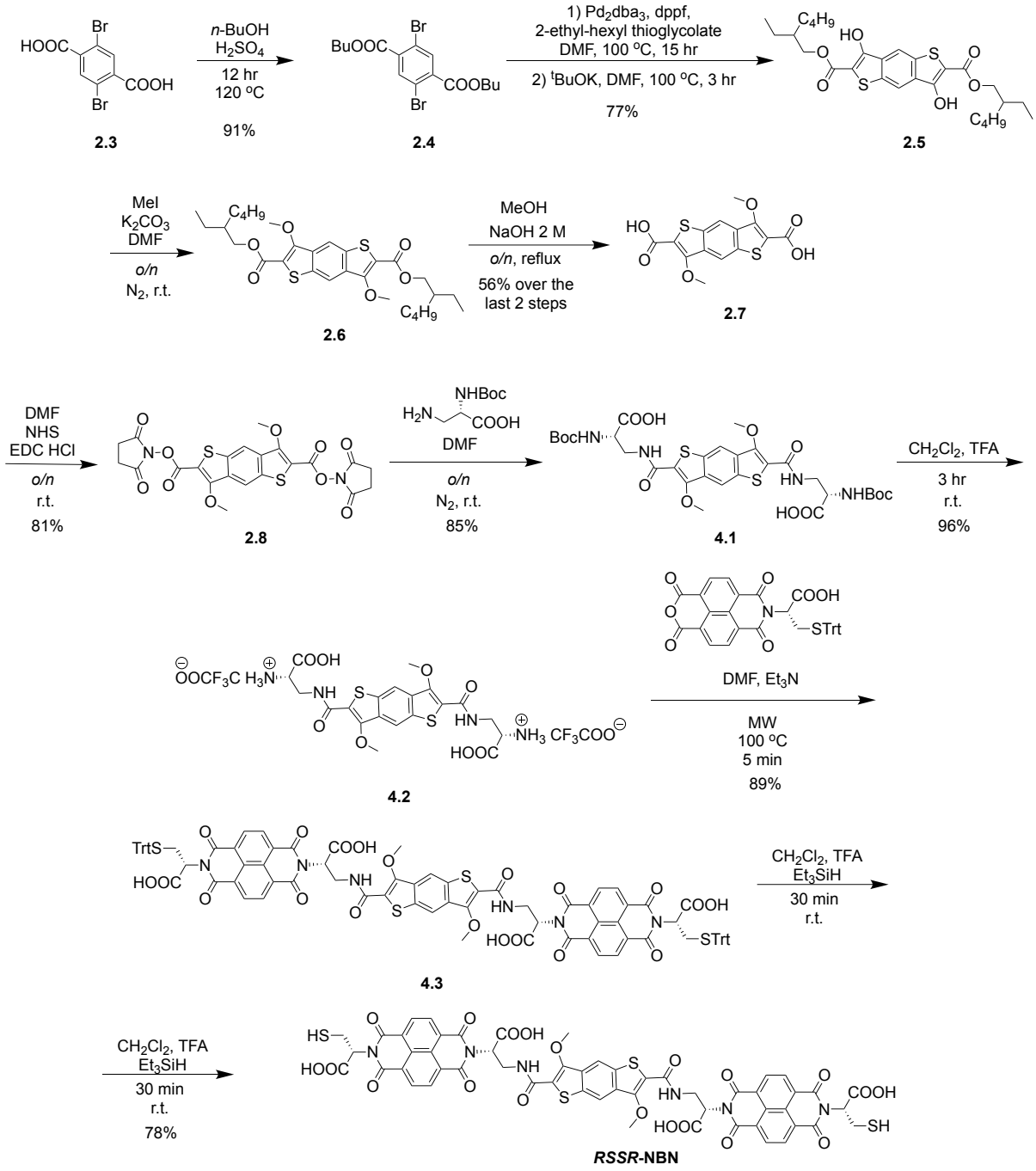
Time / min	Water (0.1% FA)	CH_3CN (0.1% FA)
0	90%	10%
20	70%	30%
35	65%	35%
37	0%	100%

Analytical HPLC method for DCLs of all libraries with iso-R,R-2 **(5 mM):** Column**:** Ultra BiPhenyl, 5 x 0.3 cm, 3 µm. Injection volume: 3 µL; Flow rate: 0.5 mL/min; Temperature: 39 °C; Run time: 43 min; Elution profile:

Method L:

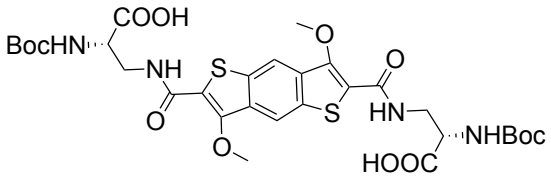
Time / min	Water (0.1% FA)	CH_3CN (0.1% FA)
0	80%	20%
5	75%	25%
20	70%	30%
30	65%	35%
35	60%	40%
40	50%	50%
43	10%	90%

Methods and synthesis for Chapter 4



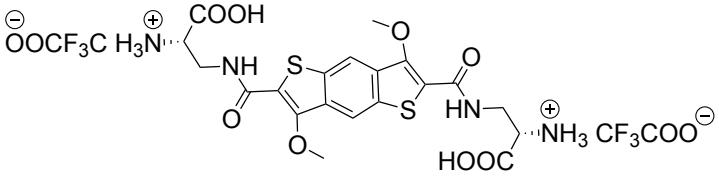
The synthesis of compounds **2.3** – **2.8** was previously described.

Synthesis of **4.1**:



In a flame-dried round-bottom flask, ***N*α-Boc-L-β-aminoalanine** (218 mg, 1.06 mmoles, 2.27 equiv) was added to a solution of **2.8** (250 mg, 0.47 mmoles, 1 equiv) in dry DMF (23 mL) under N_2 atmosphere. Dry Et_3N (0.6 mL) was added and the reaction mixture was stirred overnight at r.t. The solvent was removed under reduced pressure and a small volume of acetone was added to the mixture. This was dropwise precipitated into a vigorously stirred aqueous solution of 1 M HCl. The pale yellow precipitate was collected by filtration and dried (282.5 mg, 0.4 mmoles, 85%). ^{1}H NMR (500 MHz, DMSO-d_6): δ 12.78 (br, 2H), 8.63 (s, 2H), 8.22 (t, J = 6.0 Hz, 2H), 7.28 (d, J = 7.8 Hz, 2H), 4.17 (q, J = 7.2 Hz, 2H), 4.12 (s, 6H), 3.76 (dt, J = 12.3, 5.8 Hz, 2H), 3.51 (dt, J = 14.2, 6.6 Hz, 2H), 1.36 (s, 18H). ^{13}C NMR (125 MHz, DMSO-d_6) δ 172.7, 161.2, 155.9, 150.0, 133.9, 132.6, 124.5, 118.0, 78.8, 62.7, 53.8, 28.6. FTMS +pNSI: *m/z* calcd for $C_{30}H_{38}N_4O_{12}S_2$ $[M+H]^+$ 711.2000, found 711.2001.

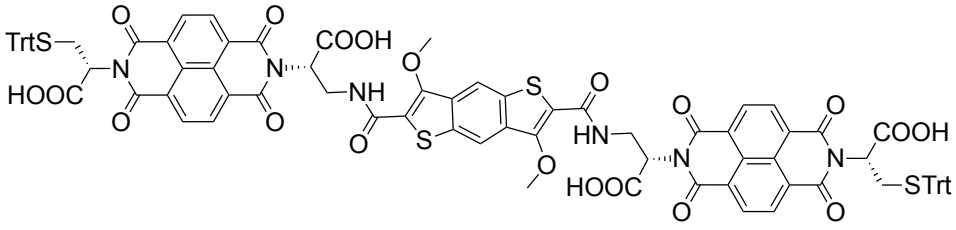
Synthesis of **4.2**:



In a round-bottom flask, **4.1** (282.5 mg, 0.4 mmoles) was dissolved in CH_2Cl_2 (2 mL) followed by addition of TFA (2 mL). The reaction mixture was stirred 3 h at r.t. and the volatiles were subsequently removed. Et_2O was added to ensure the precipitation of the desired product. The solid was filtered off and further washed with Et_2O to yield 4.2 as a red precipitate that was dried (280.2 mg, 0.38 mmoles, 96%). ^{1}H NMR (500 MHz, DMSO-d_6): δ 14.11 (br, 2H), 8.68 (s, 2H), 8.41 (br, 2H), 8.29 (br, 6H), 4.17 (s, 6H), 4.10 (br, 4H), 3.83 (dd, J = 14.3, 7.5 Hz, 4H). ^{13}C NMR (125 MHz,

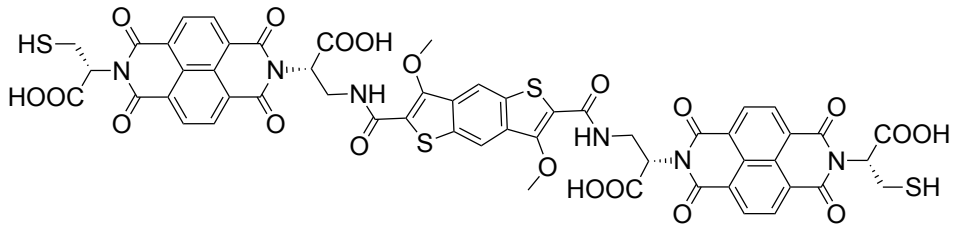
DMSO-d_6) δ 169.9, 161.9, 150.4, 134.0, 132.6, 124.3, 118.2, 62.8, 52.4. FTMS -pNSI: *m/z* calcd for $C_{20}H_{20}N_4O_8S_2$ [M-H]$^-$ 509.0806, found 509.0814.

Synthesis of **4.3**:



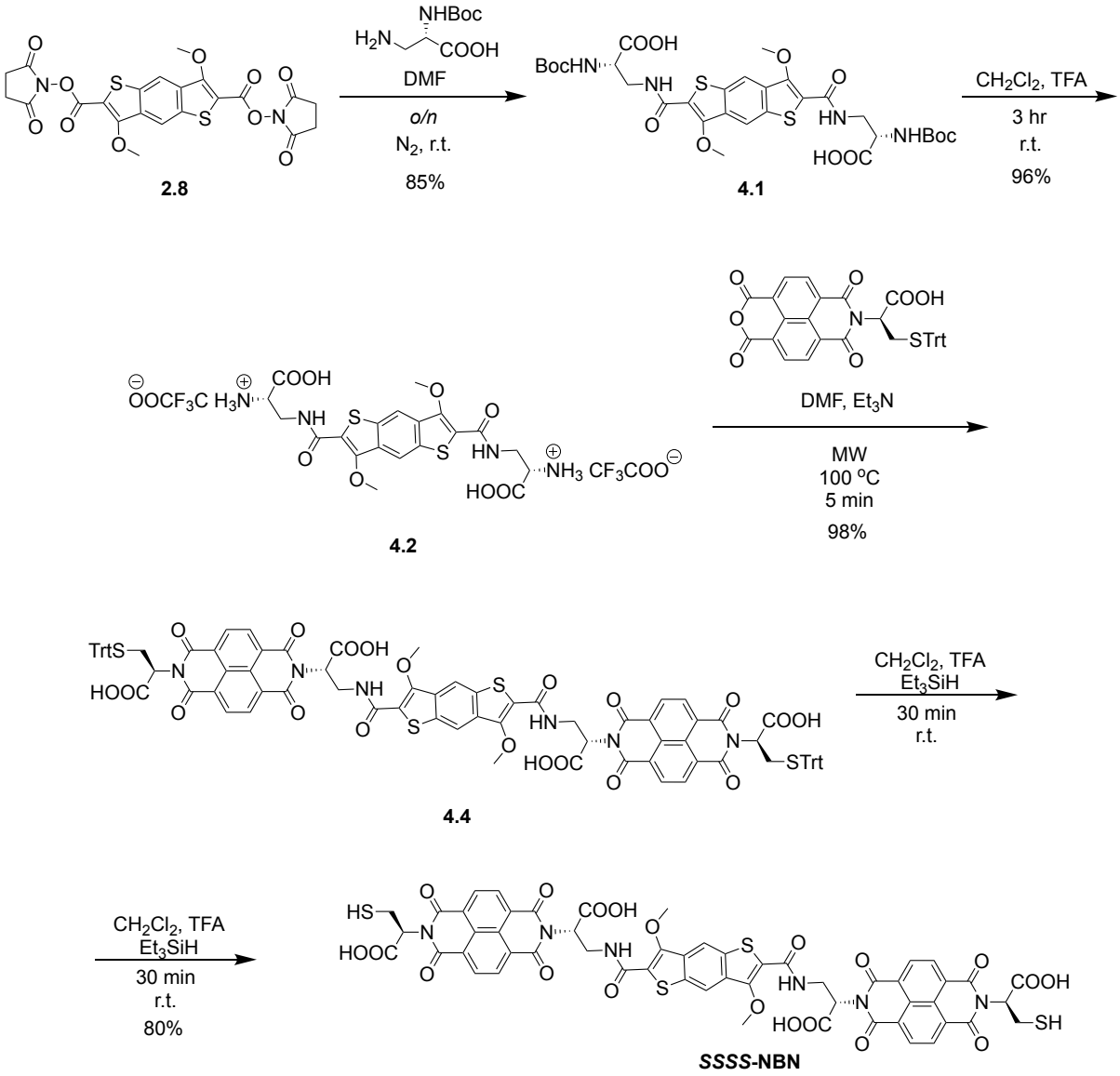
In an 8-mL microwave tube, **4.2** (100 mg, 0.14 mmoles, 1 equiv) and NMI-Trt-*R*-Cys (183.1 mg, 0.3 mmoles, 2.2 equiv) were dissolved in dry DMF (5 mL). After the addition of dry Et_3N (0.5 mL), the mixture was sonicated until no solid was present. The solution was heated under microwave irradiation at 100 °C for 5 minutes and then the solvent was removed *in vacuo*. The residue was suspended in minimum amount of acetone, and the suspension was added dropwise to a vigorously stirred aqueous solution of 1 M HCl. The red precipitate formed was filtered off and dried (205.7 mg, 0.12 mmoles, 89%). ^{1}H NMR (500 MHz, DMSO-d_6): δ 13.18 (br, 4H), 8.70 (t, *J* = 9.2 Hz, 8H), 8.39 (brs, 4H – BDT + NH), 7.24 – 7.09 (m, 30H), 5.82 (dd, *J* = 9.1, 4.2 Hz, 2H), 5.53 (dd, *J* = 10.3, 4.3 Hz, 2H), 4.27 – 4.15 (m, 2H), 4.02 (dd, *J* = 15.3, 7.6 Hz, 2H), 3.89 (s, 6H), 3.18 – 3.07 (m, 2H). ^{13}C NMR (125 MHz, DMSO-d_6) δ 169.8, 169.6, 163.0, 162.4, 161.6, 149.8, 144.5, 133.7, 132.5, 132.0, 131.5, 129.4, 128.5, 127.2, 125.9, 123.9, 117.7, 67.0, 62.4, 53.6, 52.7, 30.7. FTMS -pNSI: *m/z* calcd for $C_{92}H_{64}N_6O_{20}S_4$ [M-H]$^-$ 1699.2985, found 1699.2988.

Synthesis of ***RSSR*-NBN**:



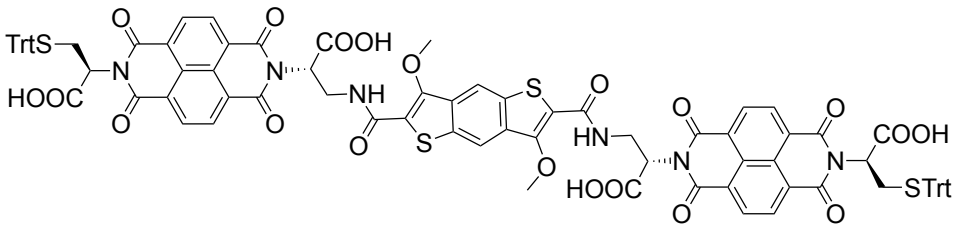
In a round-bottom flask, **4.3** (155 mg, 0.11 mmoles) was dissolved in CH_2Cl_2 (4 mL) followed by addition of TFA (4 mL) and $SiEt_3H$ (0.4 mL). The reaction mixture was

stirred 30 minutes at r.t. and the volatiles were subsequently removed. Et_2O was added to ensure the precipitation of ***RSSR*-NBN** as a red precipitate that was dried (108.1 mg, 0.09 mmoles, 78%). ^{1}H NMR (500 MHz, DMSO-d_6): δ 13.21 (br, 4H), 8.72 (t, 8.1 Hz, 8H), 8.44 (s, 2H), 8.40 (t, *J* = 6.2 Hz, 2H), 5.82 (dd, *J* = 9.2, 4.1 Hz, 2H), 5.70 (dd, *J* = 9.1, 5.5 Hz, 2H), 4.27 – 4.15 (m, 2H), 4.05 (dt, *J* = 14.3, 7.2 Hz, 2H), 3.90 (s, 6H), 3.18 (dt, *J* = 14.1, 9.1 Hz, 2H), 2.72 (t, *J* = 8.9 Hz, 3H). The missing β protons are under the water peak. ^{13}C NMR (125 MHz, DMSO-d_6) δ 174.7, 167.7, 167.6, 166.4, 138.4, 137.3, 131.5, 131.4, 131.3, 128.6, 110.1, 67.2, 61.3, 28.0 FTMS -pNSI: *m/z* calcd for $C_{54}H_{36}N_6O_{20}S_4$ $[M-2H]^{2-}$ 607.0361, found 607.0362.



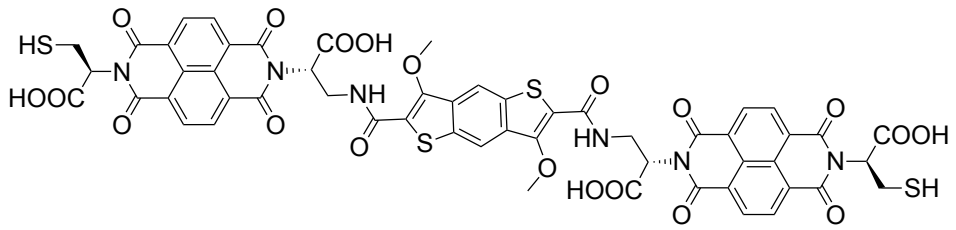
The synthesis of compounds **4.1** – **4.2** were previously described.

Synthesis of **4.4**:



In an 8-mL microwave tube, **4.2** (100.5 mg, 0.14 mmoles, 1 equiv) and NMI-Trt-*S*-Cys (183.3 mg, 0.3 mmoles, 2.2 equiv) were dissolved in dry DMF (5 mL). After the addition of dry Et_3N (0.5 mL), the mixture was sonicated until no solid was present. The solution was heated under microwave irradiation at 100 °C for 5 minutes and then the solvent was removed *in vacuo*. The residue was suspended in minimum amount of acetone, and the suspension was added dropwise to a vigorously stirred aqueous solution of 1 M HCl. The red precipitate formed was filtered off and dried (227.2 mg, 0.13 mmoles, 98%). ^{1}H NMR (500 MHz, DMSO-d_6): δ 13.18 (br, 4H), 8.71 (m, 8H), 8.40 (m, 4H), 7.23 – 7.09 (m, 30H), 5.83 (dd, J = 9.1, 4.2 Hz, 2H), 5.53 (dd, J = 10.1, 4.7 Hz, 2H), 4.28 – 4.16 (m, 2H), 4.04 (d, J = 9.2 Hz, 2H), 3.89 (s, 6H), 3.13 (dd, J = 13.1, 4.4 Hz, 2H), 3.00 – 2.81 (m, 2H). ^{13}C NMR (125 MHz, DMSO-d_6) δ 169.8, 169.6, 162.9, 162.36, 161.7, 149.9, 144.5, 133.7, 132.5, 132.0, 131.5, 129.4, 128.5, 127.2, 126.8, 125.9, 67.0, 62.5, 52.8, 30.7. FTMS -pNSI: *m/z* calcd for $C_{92}H_{64}N_6O_{20}S_4$ $[M-2H]^{2-}$ 849.1456, found 849.1459.

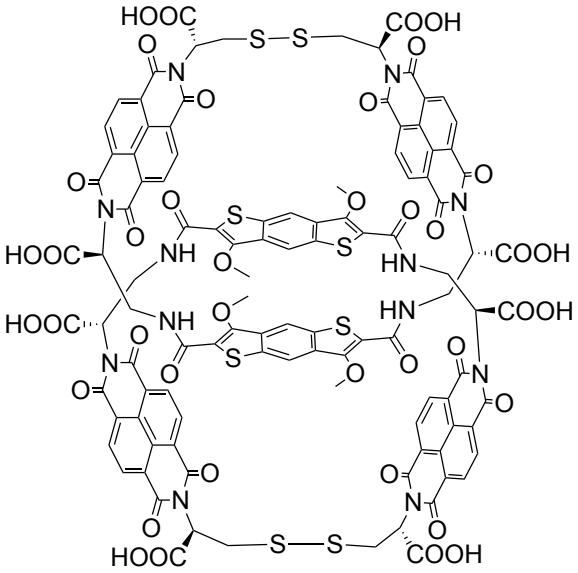
Synthesis of ***SSSS*-NBN**:



In a round-bottom flask, **4.4** (155.5 mg, 0.12 mmoles) was dissolved in CH_2Cl_2 (4 mL) followed by addition of TFA (4 mL) and $SiEt_3H$ (0.4 mL). The reaction mixture was stirred 30 minutes at r.t. and the volatiles were subsequently removed. Et_2O was added to ensure the precipitation of ***SSSS*-NBN** as a red solid that was dried (113.4 mg, 0.09 mmoles, 80%). ^{1}H NMR (500 MHz, DMSO-d_6): δ 13.21 (br, 4H), 8.72 (t, J =

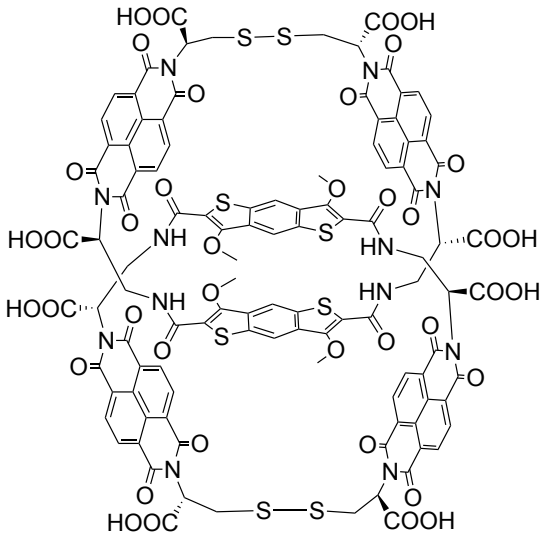
8.3 Hz, 8H), 8.44 (s, 2H), 8.40 (t, *J* = 6.3 Hz, 2H), 5.82 (dd, *J* = 9.2, 4.0 Hz, 2H), 5.70 (dd, *J* = 9.1, 5.5 Hz, 2H), 4.27 – 4.17 (m, 2H), 4.05 (dt, *J* = 14.0, 7.0 Hz, 2H), 3.90 (s, 6H), 3.19 (dt, *J* = 14.2, 9.2 Hz, 2H), 2.72 (t, *J* = 8.9 Hz, 2H). The missing β protons are under the water peak. ^{13}C NMR (125 MHz, DMSO-d_6) δ 174.7, 174.6, 167.7, 167.6, 166.4, 154.6, 138.4, 137.3, 131.5, 131.4, 131.1, 128.6, 67.2, 61.3, 58.4, 28.0. FTMS -pNSI: *m/z* calcd for $C_{54}H_{36}N_6O_{20}S_4$ $[M-2H]^{2-}$ 607.0361, found 607.0361.

***RSSR*-NBN [2]catenane** characterisation**:**



^{1}H NMR (500 MHz, D_2O): δ 8.93 (d, *J* = 7.7 Hz, 4H - NDI), 8.84 (d, *J* = 7.6 Hz, 4H - NDI), 8.76 (d, *J* = 7.7 Hz, 4H - NDI), 8.71 (d, *J* = 7.6 Hz, 4H - NDI), 6.00 (dd, *J* = 11.7, 5.2 Hz, 4H – α NDI protons), 5.70 (s, 4H – BDT protons), 5.50 (dd, *J* = 10.3, 5.5 Hz, 4H – α NDI protons), 4.55 (t, *J* = 12.7 Hz, 4H – β NDI protons), 4.00 (dd, *J* = 13.6, 5.2 Hz, 4H – β NDI protons), 3.67 (s, 12H), 3.29 (dd, *J* = 13.0, 5.5 Hz, 4H – β NDI protons), 3.07 (dd, *J* = 12.9, 10.5 Hz, 4H – β NDI protons). ^{13}C NMR (from HSQC, D_2O): δ 131.9, 131.7, 131.5, 130.8, 113.8, 61.7, 54.1, 53.6, 38.3, 34.5.

***SSSS*-NBN [2]catenane** characterisation**:**



^{1}H NMR (500 MHz, D_2O): δ 8.84 (s, 8H - NDI), 8.76 – 8.61 (m, 8H - NDI), 6.20 (s, 4H – BDT protons), 5.98 – 5.86 (m, 4H – α NDI protons), 5.73 – 5.61 (m, 4H – α NDI protons), 4.42 (t, *J* = 14.5 Hz, 4H – β NDI protons), 3.99 – 3.91 (m, 4H – β NDI protons), 3.86 (s, 12H), 3.51 (d, *J* = 14.9 Hz, 4H – β NDI protons), 3.18 (t, *J* = 13.5 Hz, 4H – β NDI protons). ^{13}C NMR (from HSQC, D_2O): δ 131.4, 131.2, 114.2, 61.8, 54.4, 52.6, 38.5, 31.6.

HPLC methods:

Analytical HPLC method for DCLs of *RSSR*-NBN and *SSSS*-NBN (for all library concentration) – for characterization of the library compounds: Column**:** Develosil C_{30}, 5 x 0.3 cm, 3 µm. Injection volume: depending on the library concentration (between 1 and 10 µL); Flow rate: 1 mL/min; Temperature: 39 °C; Run time: 14 min; Elution profile:

Method M:

Time / min	Water (0.1% Formic Acid - FA)	CH_3CN (0.1% FA)
0	70%	30%
14	45%	55%

Analytical HPLC method for DCLs of *RSSR*-NBN (0.5 mM) and *SSSS*-NBN (5 mM) – for separation purposes: Column**:** Develosil C_{30}, 5 x 0.3 cm, 3 µm. Injection volume: 10 µL (for ***RSSR*-NBN**) and 1 µL (for ***SSSS*-NBN**); Flow rate: 0.7 mL/min; Temperature: 39 °C; Run time: 20 min; Elution profile:

Method N:

Time / min	Water (0.1% Trifluoroacetic acid - TFA)	CH_3CN (0.1% TFA)
0	70%	50%
20	50%	50%

Preparative HPLC method for DCLs of *RSSR*-NBN (0.5 mM): Column**:** Develosil C_{30}, 5 x 2 cm, 5 µm; Injection volume: 950 µL; Flow rate: 5 mL/min; Temperature: 39 °C; Run time: 19 min; Elution profile:

Method O:

Time / min	Water (0.1% TFA)	CH_3CN (0.1% TFA)
0	67%	33%
17	63.5%	36.5%
18	62.5%	37.5%
19	0%	100%

Preparative HPLC method for DCLs of *RSSR*-NBN (5 mM): Column**:** Develosil C_{30}, 5 x 2 cm, 5 µm; Injection volume: 400 µL; Flow rate: 5.3 mL/min; Temperature: 39 °C; Run time: 26 min; Elution profile:

Method P:

Time / min	Water (0.1% TFA)	CH_3CN (0.1% TFA)
0	70%	30%
25	65%	35%
26	0%	100%

CD / UV-Vis studies for Chapter 4: The experiments were performed in a 1, 2 or 10 mm pathlength cuvettes depending on the concentration, with a 4 nm bandwidth. Scan mode: 1 point/nm, time-per-point 0.5 s. The VT experiments (10 °C increment and setting time 300 s) were performed in the same cuvette using a stirrer bar.

Method for CD prediction:

Each structure (3_1, 4_1, 5_1 and 5_2) was first drawn in ChemDraw®, followed by geometry optimisation in Avogadro (Force field: UFF, Conjugate gradients) and Mopac PM7 optimisation in Gabedit 2.5.0 (Solvent: Water). At the end, each structure was subjected to molecular dynamics (Amber potential: Molecular Dynamics Conformational Search, Polak-Ribiere method, step size 0.01) using one of the methods (a, b or c). The structures obtained were optimised again using PM7 and the CD spectra were predicted using ArgusLab 4.0.1 (ZINDO-RPA, 80 excitations).

Method	a	b	c
No. of geometries	20	10	10
Heating (ps)	1	1	1
Production (ps)	20	10	10
Step size (fs)	1	1	1
Heating T (K)	298	353	293
Production T (K)	298	353	313
Collide (ps^{-1})	20	20	20
Friction (ps^{-1})	40	40	40

The closeness to fit was done using the following formula:

$$r = \frac{1}{n}\sum \frac{(x_{exp}-\overline{x_{exp}}) \cdot (x_{mod}-\overline{x_{mod}})}{s_{exp} \cdot s_{mod}}$$

Where

r – closeness to fit value

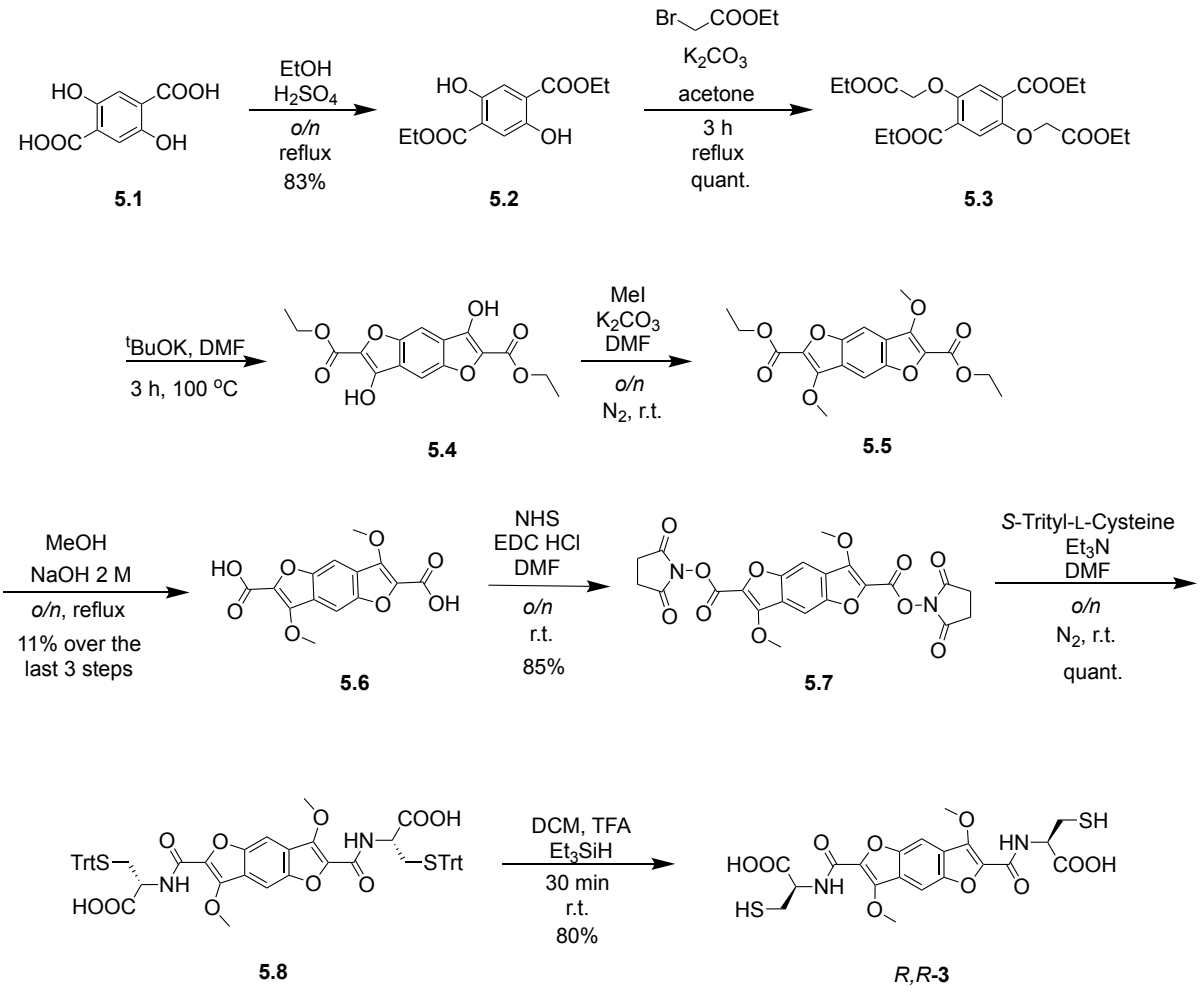
x_{exp}, x_{mod} – the experimental / theoretical value at a specific wavelength

$\overline{x_{exp}}$, $\overline{x_{mod}}$ – the average value of experimental / theoretical data

s_{exp}, s_{mod} – standard deviation of experimental / theoretical data

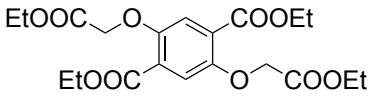
n – number of data points x_{exp}

Methods and synthesis for Chapter 5



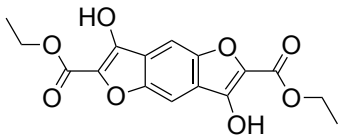
Compound **5.2** was previously synthesised.[11]

Synthesis of **5.3**:



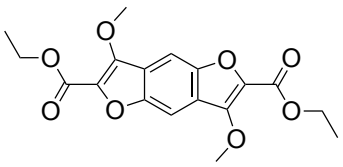
In a round-bottom flask, **5.2** (2.9 g, 11.4 mmoles, 1 equiv) was dissolved in acetone (40 mL). K_2CO_3 (19.3 g, 140 mmoles, 12.3 equiv) and ethyl bromoacetate (2.8 mL, 25.3 mmoles, 2.2 equiv) were added and the mixture was refluxed for 3 h. K_2CO_3 was filtered off and the solvent was removed under reduced pressure to yield the desired compound, which was dried (4.85 g, 11.38 mmoles, quant.). ^{1}H NMR (500 MHz, $CDCl_3$): δ 7.41 (s, 2H), 4.70 (s, 4H), 4.40 (q, *J* = 7.1 Hz, 4H), 4.29 (q, *J* = 7.1 Hz, 4H), 1.41 (t, *J* = 7.2 Hz, 6H), 1.32 (t, *J* = 7.1 Hz, 6H). ^{13}C NMR (125 MHz, $CDCl_3$) δ 168.3, 164.8, 151.8, 125.8, 118.6, 67.7, 61.6, 61.4, 14.2, 14.1. Q-TOF + ESI: *m/z* calcd for $C_{20}H_{26}O_{10}$ $[M+H]^+$ 427.1599, found 427.1600.

Synthesis of **5.4**:



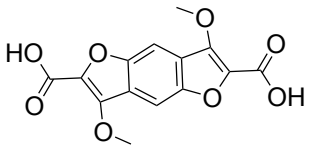
In a flame-dried round bottom flask and under N_2 atmosphere, **5.3** (4.85 g, 11.38 mmoles, 1 equiv) was dissolved in dry DMF (184 mL). After the addition of potassium *tert*-butoxide (3.83 g, 34.1 mmoles, 3 equiv), the mixture was heated for 3 h at 100 °C. For the next step, the reaction was continued by adding MeI (2.9 mL, 46.6 mmoles, 4.1 equiv), and the reaction mixture was stirred overnight at r.t. The characterisation shown below was done on a different batch of isolated compound (**5.4**). ^{1}H NMR (500 MHz, DMSO-d_6): δ 10.92 (s, 2H), 7.98 (s, 2H), 4.33 (q, *J* = 7.1 Hz, 4H), 1.33 (t, *J* = 7.1 Hz, 6H). ^{13}C NMR (125 MHz, DMSO-d_6) δ 159.6, 149.1, 147.4, 128.9, 123.8, 103.5, 60.6, 14.8. Q-TOF - ESI: *m/z* calcd for $C_{16}H_{14}O_8$ $[M-H]^-$ 333.0616, found 333.0617.

Synthesis of **5.5**:



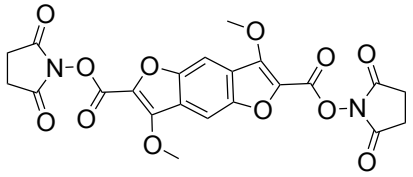
Directly to the residue synthesised above (**5.4**), water was added, and the compound was extracted with CH_2Cl_2. The organic phase was dried over anhydrous $MgSO_4$ and the solvent removed *in vacuo* to yield an oily liquid, which was dried under high pressure. At this point, the compound was impure but used as it was; no yield was determined at this stage. The presence of impurities made ^1H NMR or ^{13}C NMR assignment impossible. Q-TOF + ESI: *m/z* calcd for $C_{18}H_{18}O_8$ $[M+H]^+$ 363.1074, found 363.1079.

Synthesis of **5.6**:



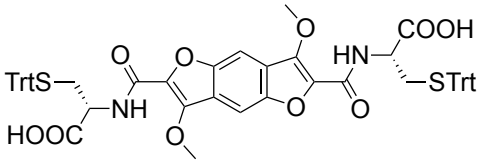
In a round-bottom flask, **5.5** (all the compound previously obtained) was dissolved in acetone (5 mL), followed by addition of MeOH (80 mL) and 2 M NaOH (80 mL). The reaction mixture was stirred overnight at reflux. Upon cooling, concentrated HCl (32%) was added until the pH became acidic and a red solid precipitated. The precipitate was filtered off and washed with CH_2Cl_2 and acetone. The precipitate was collected and dried to yield **5.6** (365.2 mg, 1.2 mmoles, 11% over 3 steps). ^1H NMR (500 MHz, DMSO-d_6): δ 13.43 (br, 2H), 8.13 (s, 2H), 4.20 (s, 6H). ^{13}C NMR (125 MHz, DMSO-d_6) δ 160.2, 149.2, 148.6, 133.8, 123.6, 104.1, 62.0. Q-TOF - ESI: *m/z* calcd for $C_{14}H_{10}O_8$ $[M-H]^-$ 305.0303, found 305.0303.

Synthesis of **5.7**:



In a round-bottom flask, **5.6** (362 g, 1.2 mmoles, 1 equiv) and *N*-hydroxysuccinimide (575.2 g, 5 mmoles, 4.2 equiv) were dissolved in dry DMF (45 mL) and cooled to 0 °C with an ice bath. EDC·HCl (953.9 g, 5 mmoles, 4.2 equiv) was added and the reaction was stirred for 15 minutes in the melting ice bath. The ice bath was removed, and the reaction mixture was further stirred overnight at r.t. The solvent was removed under reduced pressure and a small amount of acetone was added. The suspension formed was added dropwise to a vigorously stirred aqueous solution of 1 M HCl. The precipitate obtained was collected by vacuum filtration and dried under reduced pressure to obtain a brownish solid (502.6 g, 0.85 mmoles, 85%). ^{1}H NMR (500 MHz, DMSO-d_6): δ 8.57 (s, 2H), 4.42 (s, 6H), 2.90 (s, 8H). ^{13}C NMR (125 MHz, DMSO-d_6) δ 170.7, 154.7, 153.0, 150.7, 127.5, 123.0, 106.7, 62.0, 26.0. Q-TOF + ESI: *m/z* calcd for $C_{22}H_{16}N_2O_{12}$ $[M+H]^+$ 501.0776, found 501.0773.

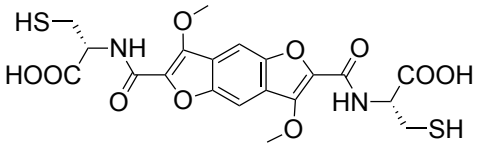
Synthesis of **5.8**:



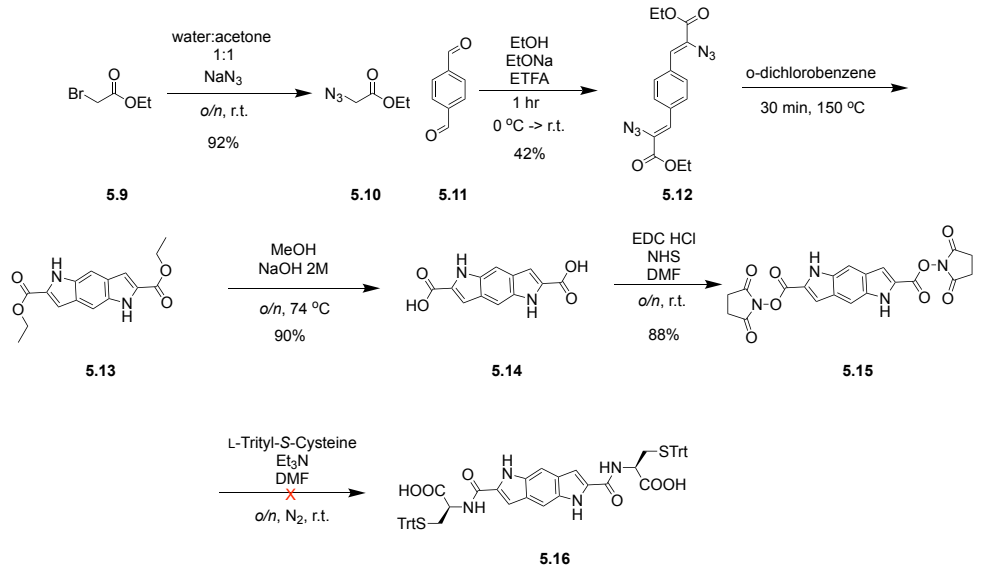
In a flame-dried round-bottom flask, *S*-trityl-L-cysteine (662.8 mg, 1.8 mmoles, 2.3 equiv) was added to a solution of **5.7** (401.8 mg, 0.8 mmoles, 1 equiv) in dry DMF (40 mL) under N_2 atmosphere. Dry Et_3N (1 mL) was added and the reaction mixture was stirred overnight at r.t. The solvent was removed under reduced pressure and a small volume of acetone was added to the mixture. This was added dropwise into a vigorously stirred aqueous solution of 1 M HCl. The reddish precipitate was collected by filtration and dried (767.5 mg, 0.77 mmoles, quant.). ^{1}H NMR (500 MHz, DMSO-d_6): δ 13.07 (br, 2H), 8.24 (d, J = 7.8 Hz, 2H), 8.16 (s, 2H), 7.35 – 7.24 (m, 30H), 4.48 (td,

J = 7.8, 4.6 Hz, 2H), 4.25 (s, 6H), 2.83 (dd, *J* = 12.3, 7.9 Hz, 2H), 2.60 (dd, *J* = 12.3, 4.6 Hz, 2H). ^{13}C NMR (125 MHz, DMSO-d_6) δ 171.7, 157.9, 148.9, 145.6, 144.6, 135.3, 129.5, 128.6, 127.3, 122.8, 104.0, 66.7, 62.0, 51.6, 33.5. Q-TOF - ESI: *m/z* calcd for $C_{58}H_{48}N_2O_{10}S_2$ [M-H]$^-$ 995.2678, found 995.2676.

Synthesis of *R,R*-**3**:

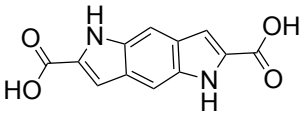


In a round-bottom flask, **5.8** (601.1 mg, 0.6 mmoles) was dissolved in CH_2Cl_2 (10 mL) followed by addition of TFA (10 mL) and $SiEt_3H$ (1 mL). The reaction mixture was stirred 50 minutes at r.t. and the volatiles were subsequently removed. Et_2O was added to ensure the precipitation of the desired product, which was filtered off and collected. The Et_2O solution was evaporated and petroleum ether was added, which was further removed by pipetting until about 5 mL remained. Et_2O was added and a precipitate was formed, which was collected by filtration. Both precipitates were confirmed to be *R,R*-**3** as a pale red precipitate (246.3 mg, 0.48 mmoles, 80%). ^{1}H NMR (500 MHz, DMSO-d_6): δ 13.19 (br, 2H), 8.20 (d, *J* = 7.8 Hz, 2H), 8.18 (s, 2H), 4.71 (td, *J* = 6.8, 4.5 Hz, 2H), 4.27 (s, 6H), 3.05 (dddd, *J* = 28.6, 14.0, 8.7, 5.0 Hz, 4H). The SH proton is under the DMSO peak. ^{13}C NMR (125 MHz, DMSO-d_6) δ 171.7, 158.1, 148.9, 145.6, 135.4, 122.7, 104.0, 62.0, 54.3, 26.0. Q-TOF - ESI: *m/z* calcd for $C_{20}H_{20}N_2O_{10}S_2$ [M-H]$^-$ 511.0487, found 511.0488.



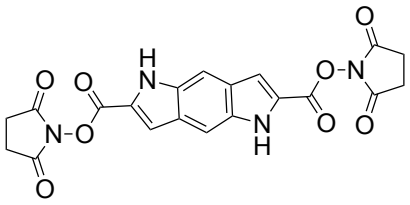
Compounds **5.9** – **5.13** were previously synthesised.[12]

Synthesis of **5.14**:



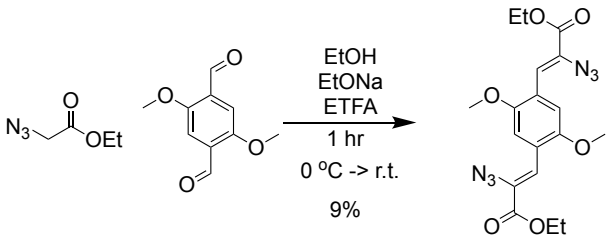
In a round-bottom flask, **5.13** (79.7 mg, 0.27 mmoles), MeOH (3 mL) and 2 M NaOH (1.5 mL) were added; the reaction mixture was stirred overnight at reflux. Upon cooling, concentrated HCl (32%) was added until the pH became acidic and a red solid precipitated. The precipitate was collected and dried to yield **5.14** (41.8 mg, 0.17 mmoles, 65%). ^{1}H NMR (500 MHz, DMSO-d_6): δ 12.85 (br, 2H), 11.44 (s, 2H), 7.28 (s, 2H), 7.14 (d, J = 2.2 Hz, 2H).

Synthesis of **5.15**:



In a round-bottom flask, **5.14** (33 mg, 0.14 mmoles, 1 equiv) and *N*-hydroxysuccinimide (68.4 mg, 0.6 mmoles, 4.4 equiv) were dissolved in dry DMF (5 mL) and cooled to 0 °C with an ice bath. EDC·HCl (104 mg, 0.54 mmoles, 4 equiv)

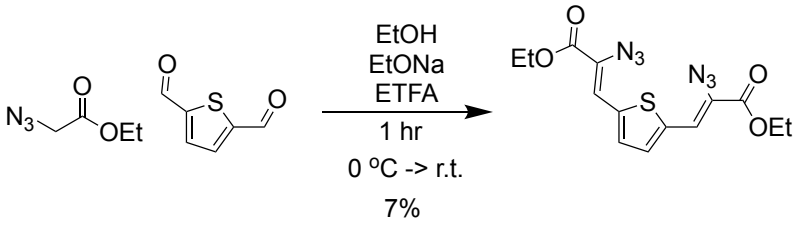
was added and the reaction mixture was stirred for 15 minutes in the melting ice bath. The ice bath was removed and the reaction mixture was further stirred overnight at r.t. The solvent was removed under reduced pressure and a small amount of acetone was added. The suspension formed was added dropwise into vigorously stirred aqueous solution of a 1 M HCl. The precipitate obtained was collected by vacuum filtration and dried under reduced pressure to obtain a brownish precipitate (44 mg, 0.1 mmoles, 74%). ^{1}H NMR (500 MHz, DMSO-d_6): δ 11.95 (s, 2H), 7.61 (s, 2H), 7.43 (s, 2H), 2.87 (s, 8H).



5.10 **5.18** **5.19**

Synthesis of **5.19**:

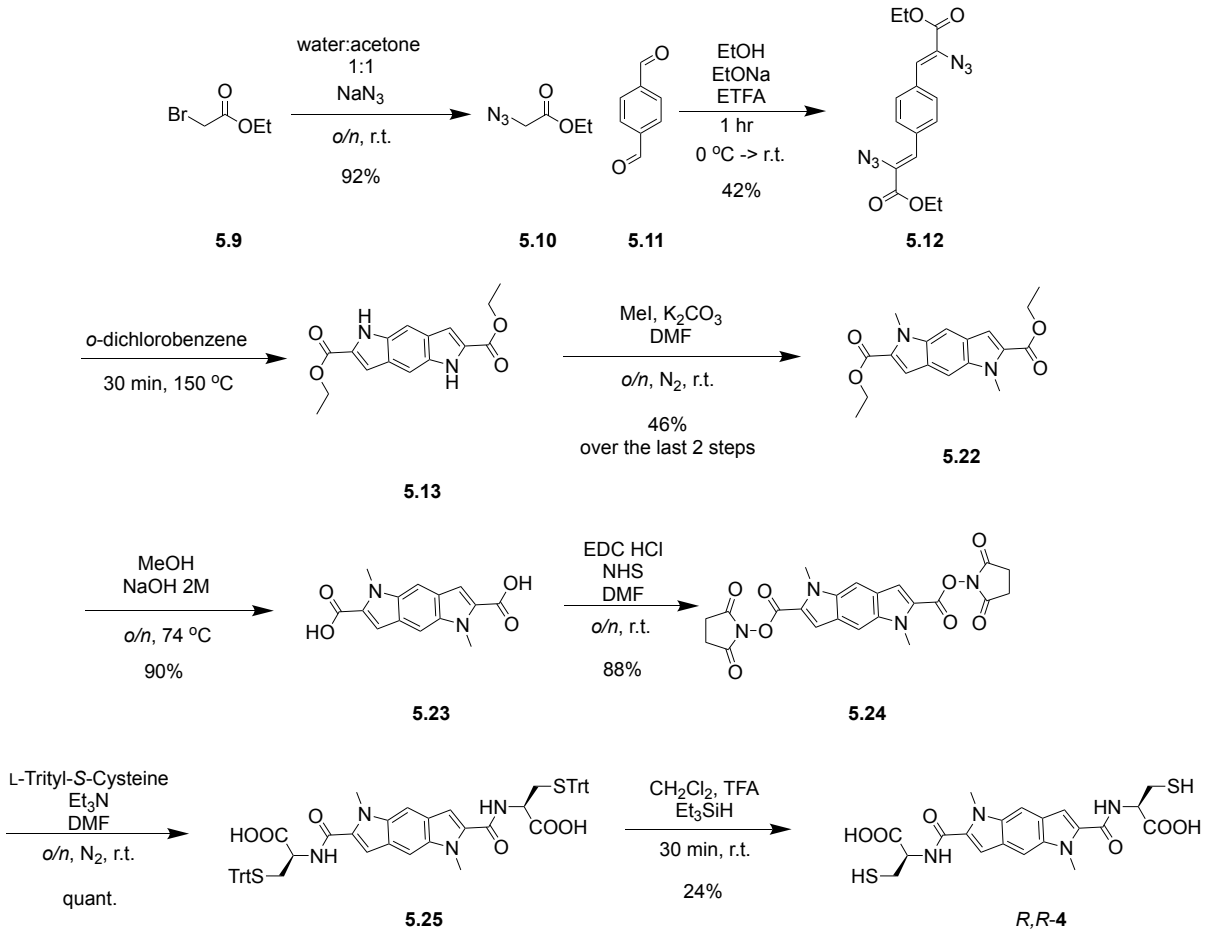
In a flame-dried round-bottom flask, a fresh solution of EtONa (10 mL, 0.25 M, 4.8 equiv) was prepared under N_2 atmosphere and cooled to 0 °C with an ice bath. In a separate round-bottom flask, ethyl trifluoroacetate (0.19 mL, 1.6 mmoles, 3.1 equiv), **5.10** (201.9 mg, 1.56 mmoles, 3 equiv), **5.18** (100 mg, 0.5 mmoles, 1 equiv) and ethanol (2.2 mL) were added. The resulting solution was added into the first one under stirring in a single portion and the ice bath was removed. After one hour, the reaction was quenched with aqueous $NH_4Cl_{(sat)}$. The precipitate formed was filtered off and kept under N_2 atmosphere (19.3 mg, 0.046 mmoles, 9%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.84 (s, 2H), 7.38 (s, 2H), 4.38 (q, *J* = 7.1 Hz, 4H), 3.89 (s, 6H), 1.41 (t, *J* = 7.1 Hz, 6H). ^{13}C NMR (125 MHz, $CDCl_3$) δ 163.6, 151.6, 125.9, 124.0, 118.3, 112.3, 62.3, 56.3, 14.2.



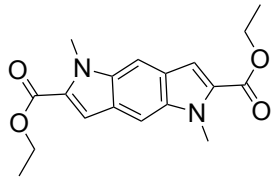
5.10 **5.20** **5.21**

Synthesis of **5.21**:

In a flame-dried round-bottom flask, a fresh solution of EtONa (16 mL, 0.46 M, 4.8 equiv) was prepared under N_2 atmosphere and cooled to 0 °C. In a separate round-bottom flask, ethyl trifluoroacetate (0.51 mL, 4.28 mmoles, 3. equiv), **5.10** (571.5 mg, 4.4 mmoles, 3.1 equiv), **5.20** (199.8 mg, 1.42 mmoles, 1 equiv) and ethanol (4.5 mL) were added. The resulting solution was added into the first one under stirring in a single portion and the ice bath was removed. After one hour, the reaction was quenched with aqueous $NH_4Cl_{(sat)}$. The precipitate formed was filtered off and kept under N_2 atmosphere (37.4 mg, 0.1 mmoles, 7%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.25 (s, 2H), 7.11 (s, 2H), 4.37 (q, *J* = 7.1 Hz, 4H), 1.40 (t, *J* = 7.1 Hz, 6H). ^{13}C NMR (125 MHz, $CDCl_3$) δ 162.9, 140.7, 131.8, 123.8, 118.6, 62.3, 14.2. The sample was sent to Swansea University for MS analysis; however, the sample was decomposing in the MS instruments due to high temperatures and no meaningful data could be acquired.

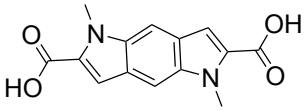


Synthesis of **5.22**:



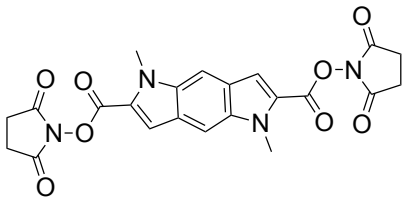
In a round-bottom flask, **5.4** (173.2 mg, 0.58 mmoles,1 equiv) was dissolved in dry DMF (35 mL) under N_2 atmosphere, followed by addition of K_2CO_3 (959.3 mg, 6.9 mmoles, 12 equiv). After stirring half an hour at r.t., MeI (0.3 mL, 4.8 mmoles, 8 equiv) was added and the reaction mixture was further stirred overnight. Water was added and the precipitate formed was filtered off. The precipitate was dried under high pressure (106.1 mg, 0.32 mmoles, 46% over the last two steps). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.44 (s, 2H), 7.34 (s, 2H), 4.39 (q, *J* = 7.1 Hz, 4H), 4.25 (s, 6H), 1.42 (t, *J* = 7.1 Hz, 6H). ^{13}C NMR (125 MHz, $CDCl_3$) δ 161.9, 129.8, 129.3, 125.8, 116.6, 114.1, 60.5, 39.0, 14.4. FTMS +pNSI: *m/z* calcd for $C_{18}H_{20}N_2O_4$ $[M+H]^+$ 329.1496, found 329.1498.

Synthesis of **5.23**:



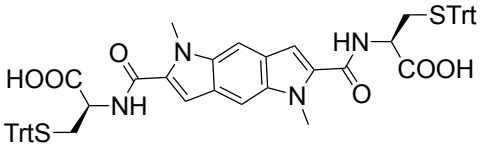
In a round-bottom flask, **5.22** (90.5 mg, 0.28 mmoles, 1 eqiuiv), MeOH (8 mL) and 2 M NaOH (4 mL) were added; the reaction mixture was stirred overnight at reflux. Upon cooling, concentrated HCl (32%) was added until the pH became acidic and the precipitate was collected and dried to yield **5.23** (67.2 mg, 0.25 mmoles, 90%). ^{1}H NMR (500 MHz, DMSO-*d*$_6$): δ 12.73 (br, 2H), 7.37 (s, 2H), 7.34 (s, 2H), 4.21 (s, 6H). ^{13}C NMR (125 MHz, DMSO-d_6) δ 163.1, 129.8, 129.4, 125.5, 116.7, 113.6, 39.1. TOF MS ASAP + (Solid): *m/z* calcd for $C_{18}H_{12}N_2O_4$ $[M+H]^+$ 273.0875, found 273.0873.

Synthesis of **5.24**:



In a round-bottom flask, **5.23** (60.3 mg, 0.22 mmoles, 1 equiv) and *N*-hydroxysuccinimide (105.5 mg, 0.91 mmoles, 4.1 equiv) were dissolved in dry DMF (7.1 mL) and cooled to 0 °C with an ice bath. EDC·HCl (176.3 mg, 0.92 mmoles, 4.2 equiv) was added and the reaction mixture was stirred for 15 minutes in the melting ice bath. The ice bath was removed and the reaction was further stirred overnight at r.t. The solvent was removed under reduced pressure and a small amount of acetone was added. The suspension formed was added dropwise into a vigorously stirred aqueous solution of 1 M HCl. The precipitate obtained was collected by vacuum filtration and dried under reduced pressure (90.2 mg, 0.19 mmoles, 88%). ^{1}H NMR (500 MHz, DMSO-d_6): δ δ 7.83 (s, 2H), 7.52 (s, 2H), 4.28 (s, 6H), 2.90 (s, 8H). ^{13}C NMR (125 MHz, DMSO-d_6) δ 171.0, 156.9, 130.5, 126.6, 122.8, 118.0, 116.7, 39.1, 26.0. FTMS +pNSI: *m/z* calcd for $C_{22}H_{18}N_4O_8$ $[M+H]^+$ 467.1197, found 467.1194.

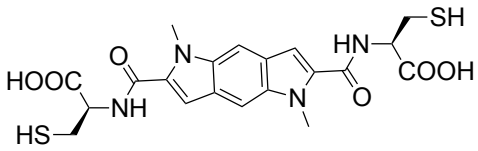
Synthesis of **5.25**:



In a flame-dried round-bottom flask, *S*-trityl-L-cysteine (105.8 mg, 0.29 mmoles, 2.3 equiv) was added to a solution of **5.25** (60.3 mg, 0.13 mmoles, 1 equiv) in dry DMF (12 mL) under N_2 atmosphere. Dry Et_3N (0.15 mL) was added and the reaction mixture was stirred overnight at r.t. The solvent was removed under reduced pressure and a small volume of acetone was added to the mixture. This was added dropwise into a vigorously stirred aqueous solution of 1 M HCl. The precipitate was collected by filtration and dried (128.6 mg, 0.13 mmoles, quant.). ^{1}H NMR (500 MHz, DMSO-d_6): δ 12.74 (br, 2H), 8.71 (d, *J* = 8.2 Hz, 2H), 7.50 – 7.02 (m, 34H), 4.33 – 4.26 (m, 2H),

4.14 (s, 6H), 2.76 (dd, *J* = 12.4, 10.0 Hz, 2H). The other β protons are under the solvent peak. ^{13}C NMR (125 MHz, DMSO-d_6) δ 172.2, 161.9, 144.7, 132.8, 129.6, 128.5, 127.3, 125.2, 116.2, 109.2, 66.7, 52.2, 39.0, 33.4, 31.1. FTMS +pNSI: *m/z* calcd for $C_{58}H_{50}N_4O_6S_2$ $[M+H]^+$ 963.3245, found 963.3250.

Synthesis of *R,R*-**4**:



In a round-bottom flask, **5.25** (100 mg, 0.1 mmoles) was dissolved in CH_2Cl_2 (3 mL) followed by addition of TFA (3 mL) and $SiEt_3H$ (0.3 mL). The reaction mixture was stirred 30 minutes at r.t. and the volatiles were subsequently removed. Et_2O was added to ensure the precipitation of *R,R*-**4** as a brownish precipitate that was dried (11.7 mg, 0.02 mmoles, 24%). ^{1}H NMR (500 MHz, DMSO-d_6): δ 12.86 (br, 2H), 8.58 (d, *J* = 8.0 Hz, 2H), 7.35 (s, 2H), 7.34 (s, 2H), 4.52 (td, *J* = 8.4, 4.4 Hz, 2H), 4.16 (s, 6H), 3.01 (ddd, *J* = 13.0, 8.2, 4.5 Hz, 2H), 2.89 (dt, *J* = 14.1, 8.4 Hz, 2H), 2.59 (t, *J* = 8.5 Hz, 2H). Q-TOF - ESI: *m/z* calcd for $C_{20}H_{22}N_4O_6S_2$ $[M-H]^+$ 477.0908, found 477.0910.

HPLC methods:

Analytical HPLC method for DCLs of all libraries of *R,R*-3 with *R,R*-1 and/or *S,S*-1 – for analytical purposes (5 mM): Column**:** Ultra BiPhenyl, 5 x 0.3 cm, 3 µm. Injection volume: 3 µL; Flow rate: 0.5 mL/min; Temperature: 39 °C; Run time: 43 min; Elution profile:

Method Q:

Time / min	Water (0.1% FA)	CH_3CN (0.1% FA)
0	80%	20%
5	75%	25%
20	70%	30%
30	65%	35%
35	60%	40%
40	50%	50%
43	10%	90%

Analytical HPLC method for DCLs of all libraries of *R,R*-3 with and without *R,R*-1 / *S,S*-1 – For analytical/semi-preparative purposes (5 mM): Column**:** Sunfire C_{18}, 15 x 0.46 cm, 5 µm. Injection volume: 5 µL; Flow rate: 1 mL/min; Temperature: 39 °C; Run time: 25 min; Elution profile:

Method R:

Time / min	Water (0.1% TFA)	CH_3CN (0.1% TFA)
0	80%	20%
5	75%	25%
25	65%	90%

Analytical HPLC method for DCLs of all libraries of *R,R*-3 with and without *R,R*-1 / *S,S*-1 – For semi-preparative purposes (5 mM): Column**:** Sunfire C_{18}, 10 x 1.0 cm, 5 µm. Injection volume: 400 µL; Flow rate: 3 mL/min; Temperature: 39 °C; Run time: 26 min; Elution profile:

Method S:

Time / min	Water (0.1% TFA)	CH_3CN (0.1% TFA)
0	80%	20%
5.25	75%	25%
25	65%	35%

Analytical HPLC method for DCLs of all libraries of *R,R*-4 with and without *R,R*-1 (5 mM): Column**:** Sunfire C_{18}, 15 x 0.46 cm, 5 µm. Injection volume: 5 µL; Flow rate: 1 mL/min; Temperature: 39 °C; Run time: 25 min; Elution profile:

Method T:

Time / min	Water (0.1% FA)	CH_3CN (0.1% FA)
0	90%	10%
25	10%	90%

CD / UV-Vis studies for Chapter 5: The experiments were performed in a 1 or 10 mm pathlength cuvettes depending on the concentration, with a 4 nm bandwidth. Scan mode: 1 point/nm, time-per-point 0.5 s. The VT experiments (10 °C increment and setting time 300 s) were performed in the same cuvette using a stirrer bar.

7.1. References

1 M. D. Hanwell, D. E. Curtis, D. C. Lonie, T. Vandermeersch, E. Zurek and G. R. Hutchison, *J. Cheminformatics*, 2012, **4**, 17.

2 J. J. P. Stewart, *J. Mol. Model.*, 2013, **19**, 1–32.

3 A.-R. Allouche, *J. Comput. Chem.*, 2011, **32**, 174–182.

4 S. Ota, S. Minami, K. Hirano, T. Satoh, Y. Ie, S. Seki, Y. Aso and M. Miura, *RSC Adv.*, 2013, **3**, 12356–12365.

5 H. Y. Au-Yeung, G. D. Pantoş and J. K. M. Sanders, *Proc. Natl. Acad. Sci.*, 2009, **106**, 10466–10470.

6 P. Pengo, G. D. Pantoş, S. Otto and J. K. M. Sanders, *J. Org. Chem.*, 2006, **71**, 7063–7066.

7 K. Tambara, N. Ponnuswamy, G. Hennrich and G. D. Pantoş, *J. Org. Chem.*, 2011, **76**, 3338–3347.

8 H. Y. Au-Yeung, G. D. Pantoş and J. K. M. Sanders, *J. Am. Chem. Soc.*, 2009, **131**, 16030–16032.

9 F. B. L. Cougnon, H. Y. Au-Yeung, G. D. Pantoş and J. K. M. Sanders, *J. Am. Chem. Soc.*, 2011, **133**, 3198–3207.

10 Y. Shi, T. Suguri, C. Dohi, H. Yamada, S. Kojima and Y. Yamamoto, *Chem. - Eur. J.*, 2013, **19**, 10672–10689.

11 X. Ma, C. Xu, J. Wang and H. Tian, *Angew. Chem. Int. Ed.*, 2018, **57**, 10854–10858.

12 W. L. H. Iv, C. S. Gelbaum, L. Gelbaum, P. Pollet, K. W. Richman, W. DuBay, J. D. Butler, G. Wells and C. L. Liotta, *RSC Adv.*, 2013, **3**, 13232–13242.